

# Analysis of Artifacts from Marshall's Blackboard Outcomes Repository

## Academic Year 2016 – 2017

**Summer Assessment Workgroup Members:** Marie Archambault, Cam Brammer, Kim DeTardo-Bora, Robert Ellison, Victor Fet, Marty Laubach, Joan St. Germain, Anita Walz, Mary Welch, Mary Beth Reynolds and Tim Melvin (Office of Assessment), and Doug Nichols (Academic Affairs Blackboard Technical Support)

### Executive Summary

#### *Background*

#### *Recommendations from the 2016 Assessment Workgroup (with current status in red)*

#### *Recommendations Concerning the General Process of Assignment Creation and Accurate Alignment to University Outcomes*

We first note that, beginning with academic year 2016-2017, faculty were asked to develop assignments that align to the outcomes as stated in Marshall University's Baccalaureate Degree Profile. In other words, we abandoned the former practice of asking instructors to indicate which performance level on the rubric they used when creating assignments. The reason for this decision was that the former rubric level descriptions were essentially different outcome statements. The Assessment Workgroup began the process of redeveloping the rubrics so that performance levels now specify *how well* each student demonstrates mastery of the university's *outcomes*, not whether or not the student achieves progressively more complex outcomes. Outlined below are concerns and recommendations from the Assessment Workgroup.

1. A major concern among the members of the Assessment Workgroup in 2016 was the large number of assignments/artifacts that the Workgroup judged to be misaligned to the outcomes/traits to which they were tagged. Several recommendations were made to improve this situation. These included:
  - Work with faculty to create assignments that align with the university outcomes addressed in Critical Thinking (CT), First Year Seminar (FYS) and Writing Intensive (WI) courses during the faculty development sessions that prepare instructors to teach these courses, as follows (**The Center for Teaching and Learning incorporates the BDP outcomes (as written) in all faculty development sessions**):
    - Center for Teaching and Learning for CT courses
    - Center for Teaching and Learning in conjunction with the Director of FYS for FYS courses
    - Center for Teaching and Learning in conjunction with the Director of Writing across the Curriculum for WI courses

- Identify model assignments from those already uploaded to GEAR and create a repository of these assignments. This repository can function as both a resource for faculty developing new assignments and a teaching tool during faculty preparation to teach the aforementioned course types. **This has not been accomplished.**
  - Ask the Center for Teaching and Learning to consider offering faculty development sessions focusing on alignment of assignments to Marshall University's outcomes. **In the Center for Teaching and Learning's (CTL) FYS and CT workshops, faculty work on creating assignments that align to the BDP outcomes. To facilitate thoughtful assignment creation, faculty have two meetings, providing an evening to reflect on how best to design the assignments to make these alignments. Also, during academic year 2016-2017, the CTL's Hedrick Faculty Teaching Fellow led a faculty learning community that developed assignments that align to *Integrative Thinking*. These assignments were shared with faculty through a series of workshops in spring 2017.**
  - Ask the General Education Council to require that all CT, INT, and MC courses include the assignment that will be used for general education assessment (i.e. GEAR upload) in course application and renewal materials and to explain explicitly how this assignment addresses the university outcome/trait to which it is aligned. **The General Education Council updated its Critical Thinking Designator and Criteria for CT course forms. The CT Designator form asks faculty to "DESCRIBE THE STUDENT PROJECT THAT WILL BE SUBMITTED USING THE ASSIGNMENT MODULE IN BLACKBOARD (which replaced GEAR in academic year 2016-2017), IDENTIFY THE BACCALAUREATE DEGREE PROFILE OUTCOME/S IT ASSESSES....."**
  - Ask that each assignment created with student artifacts uploaded into GEAR include an explicit explanation from the instructor as to how the assignment addresses the university outcome/trait(s) to which it is aligned. **Although we ask that this be done, some instructors are not including assignment instructions. We saw a rise in this practice following transition from GEAR to Blackboard.**
  - Members of the Assessment Workgroup will submit a proposal for a session to be presented at the August 2016 iPED: Inquiring Pedagogies Conference. The purpose of this session will be to overview the general education assessment process and findings, and to discuss with faculty the importance of careful assignment alignment to university outcomes. **This session was offered at the August 2016 iPED Conference.**
2. To reduce the number of artifacts from the assessment pool that must be discarded due to the Assessment Workgroup's judgment that the assignment itself does not align to the university outcome to which it was tagged, the Assessment Workgroup recommended that, in future, it evaluate each assignment for accuracy of alignment before the sample of artifacts is selected. **Due to time constraints, this was not done.**

### ***Recommendations regarding Marshall's Transition from GEAR to Blackboard Outcomes for Assessing Student Work***

Marshall began to use Blackboard Outcomes for general education assessment during academic year 2016-2017. This has some advantages over GEAR, but poses some challenges as well. Advantages include:

1. Faculty have to create an assignment and align it to university outcomes only once if the assignment and alignment is completed in their master course shell. Unless something changes (i.e. assignments are changed or updated), once alignments are made in Blackboard, they will simply be copied the next time the course is offered.

2. Faculty ask students to submit artifacts for the aligned assignment using the assignment module in Blackboard Learn. This allows the faculty member to assess the artifact for course grading purposes and the student and faculty member need do nothing else to support university assessment. For the latter purpose, Blackboard Outcomes makes a copy of the artifact (which does not include any instructor grading or comments, i.e. it is a *clean* copy) for later assessment.
3. As is the case with GEAR, when artifacts are randomly chosen for assessment in Blackboard Outcomes, course information is not available to assessors. **We were mistaken about this – course (but not instructor) information is available to assessors. We will contact Blackboard to see if this information can be eliminated.**

Blackboard Outcomes also presents challenges. These include:

1. Faculty align assignments to a university outcome and assessors use that outcome's rubric, which includes all of the outcome's traits. Because not all assignments align to every trait of the outcome, instructors have to indicate in their assignment instructions (and/or explicit explanation regarding alignment) the traits to which the assignment aligns. **Some instructors did make their trait alignments explicit; others did not.**

To help facilitate the transition from GEAR to Blackboard Outcomes, the following plans are in place.

1. Marshall's Baccalaureate Degree Profile outcomes will be entered into Blackboard. **This was accomplished.**
2. Several faculty teaching FYS, Anthropology, and Sociology courses during summer 2016 will create assignments and align them to University outcomes within Blackboard. They will use the Blackboard assignment tool and the Office of Assessment will set up artifact collection through Blackboard Outcomes. The Office of Assessment will test the Blackboard Outcomes assessment process at the end of the summer. **This was partially done. The Office of Assessment did not conduct an assessment, but did monitor that the alignments and uploads were successfully made.**
3. Fall 2016 will be a semester set aside to prepare faculty to begin using Blackboard as an artifact repository for assessment purposes. To facilitate this process, the following steps will be taken:
  - The Office of Assessment will administer a survey to all faculty teaching FYS, CT, WI, MC, INT, and SL courses. The survey will ask a series of questions that will allow us to divide the group into three cohorts (seasoned Blackboard users who routinely use the Blackboard assignment tool, Blackboard users who have not used the assignment tool, non-Blackboard users). **This was accomplished.**
  - After the survey has been completed, the Assessment Office will develop three online tutorials, one geared to each group of faculty identified above. **We worked with the Office of Academic Affairs and the MU Online Design Center to accomplish this.**
  - The Office of Assessment also will work with the Center for Teaching and Learning, the MU Design Center, and the Associate Vice President for Libraries and Online Learning to develop a schedule of training sessions for each cohort of faculty. **The MU Online Learning Design Center worked with faculty to transition from GEAR to Blackboard.**
3. During spring 2017, our hope is that all faculty teaching general education courses will begin to use Blackboard for assignment creation and student artifact collection. They will have access to the online tutorials and to training sessions as they did during the fall semester. **An**

online tutorial was created and staff from the MU-Online Design Center worked with faculty to align assignments to Marshall's BDP outcomes.

### ***Longitudinal Analysis***

For the initial assessment of artifacts uploaded to GEAR (summer 2013), all artifacts assessed were drawn from the university's First Year Seminar in Critical Thinking (FYS) course and we used these artifacts to assess all nine university outcomes. Mean performance across students ranged from a low of 0 for *Intercultural Thinking* (communication with other cultures) to a high of 1.24 for *Communication Fluency* (design/organization and diction). However, since artifacts were spread among so many outcomes, many traits had very small numbers (9 for communication with other cultures as compared to 24 for design/organization and 23 for diction). Other than the fact that all students included in the 2013 sample were freshmen, low means can be attributed to the fact that we had not yet settled on a score for misaligned artifacts, defaulting many of the scores to 0.

The second assessment of artifacts uploaded to GEAR (summer 2014) also included all nine outcomes, but we included artifacts from *Multicultural*, *International*, *Service Learning*, and *Writing Intensive* courses, in addition to those from FYS. The sample, however, continued to be skewed toward artifacts from lower level courses with freshman being the modal class rank for student artifacts in our sample. We decided to assign special codes to artifacts we felt to be misaligned to the outcomes or in cases of student upload or other technical issues that prevented assessment. This allowed us to see which outcomes/traits resulted in the greatest amount of confusion during the outcome/trait alignment process and resulted in recommendations to make sure instructors uploaded assignment instructions, specified the primary outcome to which their assignment aligned, and identified the performance level to which the assignment was written. Due to assessing all nine university outcomes again in 2014, we continued to have small numbers of artifacts aligned to each outcome, which led to the recommendation that we choose only three outcomes to assess in 2015, three more in 2016, and the last three in 2017 and continue to assess on a three-year cycle.

The third assessment of artifacts uploaded to GEAR (summer 2015) consisted of an in-depth assessment of artifacts that instructors aligned to the following outcomes as primary: *Intercultural Thinking* (due to sampling error, five of the alignments for *Intercultural Thinking* were secondary), *Ethical and Civic Thinking*, and *Communication Fluency*. One hundred eight artifacts were included for each outcome, resulting in a total of 324 artifacts. This sample resulted in higher numbers for each outcome trait. Results from summer 2015 suggested a need to redesign rubrics to be continuous, rather than categorical, in nature.

Finally, assessment data from 2013-2014 and 2014-2015 showed that Marshall's students improved their writing skills as they moved through the curriculum and, specifically, as they passed from 100/200 level writing intensive courses to 300/400 level writing intensive courses.

## ***Procedures for 2017 Assessment***

### ***General Procedures***

In summer 2017 we evaluated student artifacts produced in response to course assignments aligned to *Creative Thinking*, *Inquiry-Based Thinking*, and *Quantitative Thinking* that were uploaded to Blackboard during academic year 2016-2017. In May 2017 a group of nine faculty representing several academic colleges from across the university evaluated a sample of these artifacts using outcome specific rubrics. These rubrics which, as noted above, were revised prior to scoring, are included in the supporting documentation. Our sample initially consisted of 324 artifacts, 108 per outcome. However, during scoring we discovered that 12 artifacts (3 aligned to *Creative*, 6 to *Inquiry-Based*, and 3 to *Quantitative Thinking*) were not able to be opened or otherwise accessed by the reviewers for scoring. This reduced the number of usable artifacts to 312 (105 *Creative Thinking*, 102 *Inquiry-Based Thinking*, and 105 *Quantitative Thinking*). Reviewers further determined that 53 artifacts (24 *Creative Thinking*, 13 *Inquiry-Based Thinking*, and 16 *Quantitative Thinking*) were misaligned with all of the traits of the outcomes to which they had been tagged. This reduced the number of scorable artifacts to 259 (81 *Creative Thinking* and 89 each for *Inquiry-Based* and *Quantitative Thinking*). Each artifact was read by two independent reviewers (to arrive either at scores or to agreements of nonalignment for specific traits of each outcome). This project was coordinated by the Office of Assessment.

### ***Scoring Procedures***

Evaluators assessed each artifact using the following scale:

<b>Special Scoring Codes</b>	
<b>Score</b>	<b>Explanation</b>
N/A	In the opinion of the evaluator, the artifact was misaligned with the outcome/trait to which the instructor had tagged it.
Error	The student did not upload the correct assignment or there was a technical problem with the upload that prevented the artifact from being opened or assessed.
<b>Regular Scoring Codes</b>	
These codes were given to artifacts that, in the opinion of the evaluator, were aligned with appropriate outcomes/traits and contained enough information to allow assessment.	
0	The artifact did not demonstrate the minimum level of performance expected at the introductory level.
1	The artifact demonstrated introductory level performance.
2	The artifact demonstrated milestone level performance.
3	The artifact demonstrated capstone level performance.
4	The artifact demonstrated advanced level performance.

Please see the supporting information that follows this summary for a detailed explanation of scoring procedures.

### ***General Information about the Sample***

Two hundred forty-four (244; 75%) of the artifacts in our sample were drawn from courses at the 100/200 level, with the remaining 80 (25%) drawn from courses at the 300/400 level.

### ***Results and Analysis***

One challenge in reporting results of Blackboard assessment is that, although we assessed 324 artifacts (each of which was aligned to one BDP outcome), results were analyzed by each outcome trait. The total number of traits across the three outcomes was 11 (4 each for *Inquiry-Based* and *Quantitative Thinking*, and 3 for *Creative Thinking*). As mentioned previously, 12 artifacts were not able to be assessed due to upload or artifact file error, reducing the number of readable artifacts to 312. Of those, assessors agreed that 53 did not align to any trait of the outcome to which they were tagged. This left 259 scorable artifacts. However, not all of those artifacts aligned to every trait of the outcome to which it was tagged. A perusal of our supporting documentation shows that the artifacts evaluated by the Assessment Workgroup tagged to a total of 728 traits (206 for *Creative Thinking*, 288 for *Inquiry-Based Thinking*, and 234 for *Quantitative Thinking*), all of which were usable in calculating means. As can be seen in the chart below, the numbers were spread fairly evenly among the traits of *Creative Thinking*, fairly evenly for the traits of *Inquiry-Based Thinking* (although artifacts aligning to problem/question were fewer than those that aligned to the others); however, there were visible differences when considering the artifacts aligning to each of the traits of *Quantitative Thinking*, with few aligning to either *visual representation* or *statistics*.

Outcome	Trait	Total Traits Aligned
Creative Thinking	Ambiguities and Possibilities	69
	Risk Taking	62
	Synthesis and Innovation	75
Inquiry-Based Thinking	Problem/Question	62
	Research of Existing Knowledge	70
	Data Collection and Analysis	78
	Conclusions	78
Quantitative Thinking	Context	87
	Estimation	72
	Visual Representation	41
	Statistics	34
Totals		728

*Creative Thinking* means did not differ significantly based on course level for any trait. Students enrolled in courses at the 300/400 levels had significantly higher means for *Inquiry-Based Thinking* (*problem/question, data collection/analysis, and conclusions*) than did students enrolled in 100/200 level courses. Students enrolled in courses at the 300/400 levels had significantly higher means for *Quantitative Thinking* (*visual representation and statistics*) than did students enrolled in 100/200 level courses. However, we note that the number of *Quantitative Thinking* artifacts at the 300/400 level were small (only 8 for each of these traits).

Overall results showed mean performance for traits to range from 1.21 (*Quantitative Thinking: statistics*) to 2.0 (*Inquiry-Based Thinking: conclusions*). Mean performance for artifacts uploaded from 100/200 level courses ranged from 1.0 (*Quantitative Thinking: statistics*) to 1.92 (*Creative Thinking: risk taking*) and for 300/400 level courses from 1.6 (*Quantitative Thinking: estimation*) to 2.66 (*Inquiry-Based Thinking: conclusions*). Although there does not appear to be an overall strength for our students, their weakest performance was in *Quantitative Thinking* (*visual representation and statistics*). We note, however, that these two traits of *Quantitative Thinking* were the traits to which assignments least frequently aligned.

### ***Results for Course Type***

Analyzing results by course type posed several challenges. First, the only course type that is unique (i.e. can have only one course type attribute) is First Year Seminar in Critical Thinking (FYS). Courses can have the other attributes (Critical Thinking [CT], Multicultural [MC], International [INT], Writing Intensive [WI], Service Learning [SL], and Core II) in combination (and many do). So, when analyzing results by course type, we included all courses with the attribute we wanted to assess; this resulted in some courses being included in the analysis for more than one course type. Because the number of courses with INT and MC attributes in our sample was small, we did not conduct analyses of these course types. We also note that MC and INT courses have been asked to create assignments and ask students to upload artifacts whose primary alignment is to *Intercultural Thinking*, an outcome we did not assess this cycle. SL courses (which align to *Ethical and Civic Thinking*) were not included in our sample this year.

### ***Critical Thinking (CT) Courses***

CT courses in the assessment sample included those that aligned to each of the outcomes assessed: *Creative Thinking*, *Inquiry-Based Thinking*, and *Quantitative Thinking*. All CT courses are at the 100/200 level. Results are below:

<b>Creative Thinking</b>			<b>Inquiry-Based Thinking</b>			<b>Quantitative Thinking</b>		
<b>Trait</b>	Number	Mean Score	<b>Trait</b>	Number	Mean Score	<b>Trait</b>	Number	Mean Score
<b>Ambiguities and Possibilities</b>	14	1.36	<b>Problem/Question</b>	15	1.7	<b>Context</b>	72	1.49
<b>Risk Taking</b>	14	2.07	<b>Research of Existing Knowledge</b>	22	1.25	<b>Estimation</b>	60	1.43
<b>Synthesis and Innovation</b>	17	1.35	<b>Data Collection and Analysis</b>	29	1.76	<b>Visual Representation</b>	33	1.06
			<b>Conclusions</b>	28	1.88	<b>Statistics</b>	26	1.0

These results must be interpreted with caution; however, overall means were 1.67 for *Inquiry-Based Thinking*, 1.58 for *Creative Thinking*, and 1.33 for *Quantitative Thinking*.

### ***Core II Courses***

Core II courses in the assessment sample included those that aligned to each of the outcomes assessed: *Creative Thinking*, *Inquiry-Based Thinking*, and *Quantitative Thinking*. All Core II courses are at the 100/200 level. Results are below:

<b>Creative Thinking</b>			<b>Inquiry-Based Thinking</b>			<b>Quantitative Thinking</b>		
<b>Trait</b>	Number	Mean Score	<b>Trait</b>	Number	Mean Score	<b>Trait</b>	Number	Mean Score
<b>Ambiguities and Possibilities</b>	22	1.64	<b>Problem/Question</b>	15	1.67	<b>Context</b>	63	1.45
<b>Risk Taking</b>	18	2.33	<b>Research of Existing Knowledge</b>	24	1.44	<b>Estimation</b>	54	1.37
<b>Synthesis and Innovation</b>	25	1.98	<b>Data Collection and Analysis</b>	22	1.73	<b>Visual Representation</b>	28	1.04
			<b>Conclusions</b>	22	1.93	<b>Statistics</b>	21	0.83

These results must be interpreted with caution; however, overall means were 1.96 for *Creative Thinking*, 1.69 for *Inquiry-Based Thinking*, and 1.28 for *Quantitative Thinking*.



### ***First Year Seminar in Critical Thinking (FYS) Courses***

FYS courses in the assessment sample included those that aligned to each of the outcomes assessed: *Creative Thinking*, *Inquiry-Based Thinking*, and *Quantitative Thinking*. FYS is, by definition, at the 100 level. Results are below:

<b>Creative Thinking</b>			<b>Inquiry-Based Thinking</b>			<b>Quantitative Thinking</b>		
<b>Trait</b>	<b>Number</b>	<b>Mean Score</b>	<b>Trait</b>	<b>Number</b>	<b>Mean Score</b>	<b>Trait</b>	<b>Number</b>	<b>Mean Score</b>
<b>Ambiguities and Possibilities</b>	8	1.44	<b>Problem/Question</b>	22	1.34	<b>Context</b>	N/A	N/A
<b>Risk Taking</b>	7	1.36	<b>Research of Existing Knowledge</b>	22	1.41	<b>Estimation</b>	N/A	N/A
<b>Synthesis and Innovation</b>	8	0.88	<b>Data Collection and Analysis</b>	23	1.39	<b>Visual Representation</b>	N/A	N/A
			<b>Conclusions</b>	22	1.46	<b>Statistics</b>	N/A	N/A

These results must be interpreted with caution; however, overall means were 1.4 for *Inquiry-Based Thinking*, 1.22 for *Creative Thinking*, and there were no FYS artifacts aligned to *Quantitative Thinking*. We note that neither *Quantitative* nor *Creative Thinking* are course outcomes for FYS.

### ***Writing Intensive (WI) Courses***

WI courses in the assessment sample aligned to all outcomes assessed: *Creative Thinking*, *Inquiry-Based Thinking*, and *Quantitative Thinking*. Results are given below by course level for *Creative Thinking*:

<b>Trait</b>	<b>Course Level</b>	<b>Number</b>	<b>Mean Score</b>
<b>Ambiguities and Possibilities</b>	100/200	19	1.47
	300/400	26	1.77
<b>Risk Taking</b>	100/200	19	2.03
	300/400	24	1.69
<b>Synthesis and Innovation</b>	100/200	23	1.65
	300/400	28	1.65

These results must be interpreted with caution; however, overall means were 1.71 for *Creative Thinking* (100/200 level) and 1.7 (300/400 level). Overall performance on this outcome was the same regardless of course level.

WI results are given below by course level for *Inquiry-Based Thinking*:

Trait	Course Level	Number	Mean Score
Problem/Question	100/200	9	1.67
	300/400	16	2.34
Research of Existing Knowledge	100/200	11	1.18
	300/400	15	2.03
Data Collection and Analysis	100/200	10	1.8
	300/400	16	2.34
Conclusions	100/200	9	2.0
	300/400	17	2.68

These results must be interpreted with caution; however, overall means were 1.64 for *Inquiry-Based Thinking* (100/200 level) and 2.36 (300/400 level). Overall performance on this outcome was higher for 300/400 than for 100/200 level courses.

WI results are given below by course level for *Quantitative Thinking*:

Trait	Course Level	Number	Mean Score
Context	100/200	14	1.39
	300/400	10	1.75
Estimation	100/200	13	1.31
	300/400	8	1.5
Visual Representation	100/200	11	1.05
	300/400	6	2.0
Statistics	100/200	11	0.77
	300/400	6	1.75

These results must be interpreted with caution; however, overall means were 1.15 for *Quantitative Thinking* (100/200 level) and 1.73 (300/400 level). We note that there were fewer artifacts from 300/400 than from 100/200 level courses.

### Conclusion

The highest overall mean score was 1.77 for *Inquiry-Based Thinking*. This outcome also showed the largest growth in student performance between 100/200 and 300/400 level courses, with an overall mean of 1.57 for 100/200 level courses as compared to 2.3 for 300/400 level courses. As noted earlier, mean differences for three traits of *Inquiry-Based Thinking* (problem/question, data collection and analysis, and conclusions) were significantly higher for 300/400 than for 100/200 level courses. Even when statistical significance was not achieved, results showed that overall means trended in a higher direction for 300/400 as compared to 100/200 level courses for all outcomes (1.65 for 100/200 as

compared to 1.7 for 300/400 for *Creative Thinking* and 1.35 100/200 as compared to 1.82 for 300/400 for Quantitative Thinking. Traits aligned to least frequently were the *visual representation* and *statistics* traits of *Quantitative Thinking*.

### ***Recommendations from the 2017 Assessment Workgroup***

#### ***Recommendations Concerning the General Process of Assignment Creation and Accurate Alignment to University Outcomes***

We first note that, beginning with academic year 2016-2017, faculty were asked to develop assignments that aligned to the outcomes as stated in Marshall University's Baccalaureate Degree Profile (BDP). We abandoned the former practice of asking instructors to indicate which performance level on the rubric they used when creating assignments. The Assessment Workgroup began the process of redeveloping rubrics for each of the BDP outcomes so that performance levels now specify *how well* each student demonstrates mastery of the university's *outcomes*, not whether the student achieves progressively more complex outcomes. Outlined below are concerns and recommendations from the Assessment Workgroup.

The transition from our former General Education Assessment Repository to Blackboard for purposes of assessment is off to a good start; however, the Summer Assessment Workgroup made the following recommendations to improve faculty understanding of this process.

1. Staff from the Assessment Office and the MU Online Design Center should schedule meetings with small groups of faculty to discuss, demonstrate, and answer questions about the process of creating assignments in Blackboard's Assignment Module, aligning those assignments to one (or more) of Marshall's BDP outcomes, and having students submit their assignment artifacts using the Blackboard Assignment Module. The Workgroup recommended that this process begin with the staff requesting to be on the schedule of a Chairs' meeting and then following this up with visits to the faculty in as many departments as possible.
2. In meetings with faculty, Assessment and Design Center staff should emphasize the importance of the inclusion of assignment instructions in Blackboard that explain in some detail how the assignment addresses the BDP outcome to which the faculty member aligned it. If the assignment is meant to address some (but not all) traits of the outcome, the assignment instructions should include the traits that are addressed.
3. All assignment artifacts that students submit to the Blackboard assignment module for purposes of assessing Marshall's BDP should include *process statements* (aka reflection papers). In other words, each student should describe the *process* s/he used to complete the assignment. This reflection on the process should clearly explain how the assignment helped the student achieve the BDP outcome to which the assignment was aligned.

### ***Recommendations Concerning the Blackboard Outcomes Assessment Tool***

The following items are issues that we will ask Blackboard to address; however, we understand that Blackboard is a large company with many clients and must prioritize improvements to the product. So, while we are hopeful that many of our concerns will be addressed, we realize that addressing them all may take some time.

1. During our assessment cycle, each assessor's artifact queue disappeared upon completion of scoring. This was problematic when score disagreements between the two raters needed discussion.
2. We use an assessment process where each artifact is reviewed by two independent reviewers. The random reviewer assignment process that Blackboard uses is too simplistic. We had a total of nine reviewers and each of the nine reviewers had only two review partners for the Blackboard artifact reviews, whereas for the non-Blackboard part of our assessment process, we were able to pair each of our nine reviewers with each of the other eight people on the Assessment team.
3. Blackboard does not accommodate a third reader for those artifacts for which the original two readers cannot agree on a final score. We had to complete the third reader process this year outside of the Blackboard platform.
4. One of our team members noted that, in Blackboard Learn, course instructors can evaluate student work by having an artifact and rubric next to each other on the computer screen. This was not possible for assessors using Blackboard Outcomes.
5. Course names were visible to assessors. We would prefer that course information not be visible to assessors.
6. Artifacts did not have unique identifiers in the data download. Rather, each student had an anonymized identifier. Unfortunately, in one project, we had several students who had more than one assignment artifact in our assessment pool. While we were able to make sure we coded each assignment correctly, it took some time and checking to do this.
7. Some of the comment columns contained an excessive amount of HTML code, making the comments almost impossible to read.
8. We had several other technical questions which we will send to Blackboard.

### ***Recommendations Concerning the Potential Use of Value Rubrics Developed by the American Association of Colleges and Universities (AAC&U)***

There was discussion about the potential benefits of using rubrics created and validated by the American Association of Colleges and Universities (AAC&U). These AAC&U Value Rubrics have been tested and used widely throughout the United States. The Assessment Workgroup conducted a pilot in which they scored a very small sample of capstone project artifacts using the AAC&U's *Critical Thinking* and *Written Communication* Value rubrics. The group found these rubrics easy to use and their scoring resulted in very few scores of N/A. The Workgroup decided to extend this pilot project to next year's assessment. The pilot will work as outlined below.

1. Course instructors will continue to create assignments using the Assignment Module in Blackboard. Instructors will align the assignment to the appropriate BDP outcome (or outcomes). Students will submit assignment artifacts using the Blackboard Assignment Module.

2. Prior to time for the Summer Assessment Workgroup to begin its work in May 2018, we will again create collections for artifacts aligned to the same outcomes we assessed in May 2017 (*Creative*, *Inquiry-Based*, and *Quantitative Thinking*).
3. In May/June 2018, assessors will score each artifact with two rubrics as follows – *Creative Thinking* (Marshall's *Creative Thinking* rubric and the AAC&U *Creative Thinking* Value Rubric); *Inquiry-Based Thinking* (Marshall's *Inquiry-Based Thinking* rubric and the AAC&U *Critical Thinking* Value Rubric); *Quantitative Thinking* (Marshall's *Quantitative Thinking* rubric and the AAC&U's *Quantitative Literacy* Value Rubric). This procedure will help us to address the following issues that emerged during our discussions:
  - There was concern that, in an effort to create an outcome for *Creative Thinking* that would include all disciplines, we may have made it more difficult for programs in the traditional fine arts disciplines to create assignments that align to Marshall's outcome and rubric. Although the three traits of Marshall's *Creative Thinking* rubric had similar numbers of usable scores, we noted that more artifacts were judged not to align with the outcome at all than was the case for the other two outcomes we assessed this year.
  - For *Inquiry-Based Thinking*, there is a concern that Marshall's outcome and rubric are geared too specifically to traditional scientific fields and are not as applicable as they should be to assignments from fields in the liberal, visual, and performing arts. We believe that the AAC&U's *Critical Thinking* Value Rubric may be more applicable to all fields of study.
  - For *Quantitative Thinking*, there was concern that very few assignment artifacts aligned to two of Marshall's outcome traits (visual representation and statistics). There is a greater difference between Marshall's *Quantitative Thinking* rubric and the AAC&U's *Quantitative Literacy* Value Rubric than between the other Marshall and AAC&U cognates and using both rubrics in next year's assessment has the potential to help us determine which works better for our instructors and students.

# Supporting Documentation



**General Education  
Blackboard Artifact Assessment**

**Academic Year 2016 – 2017**

# Outcomes Assessed

Outcome	Abbreviation	Traits	Abbreviations
Creative Thinking	CRT	Ambiguities and Possibilities	A & P
		Risk Taking	Risk
		Synthesis/Innovation	Innovation
Inquiry-Based Thinking	IBT	Problem/Question	Question
		Research of Existing Knowledge	Knowledge
		Data Collection and Analysis	Analysis
		Conclusions	Conclusions
Quantitative Thinking	QT	Context	Context
		Estimation	Estimation
		Visual Representation	Visual
		Statistics	Statistics



# Course Types

Course Type	Abbreviation
First Year Seminar in Critical Thinking	FYS
Critical Thinking	CT
Multicultural	MC
International	INT
Writing Intensive	WI
Service Learning	SL (Not included in this year's assessment)
Core II	Core II

# Course Types in CRT, IBT, and QT Outcome Population

Each Course Counted Separately for Each Category

(i.e. sample  $n$  does not add to 324)

Course Type	Population $n$	Sample $n$	Percent
FYS	626	54	9%
CT	1,272	154	12%
MC	503	49	10%
INT	105	12	11%
WI	983	128	13%
Core II	1,195	150	13%
<b>Total</b>	<b>4,684</b>	<b>547</b>	<b>12%</b>

Population/Sample Comparisons for Marshall's Course Types by Course Level  
 Each Course Counted Separately for Each Category  
 (i.e. sample *n* does not add to 324)

Course Type	Course Level = 100/200			Course Level = 300/400		
	Population	Sample	Percent	Population	Sample	Percent
FYS	626	54	9%	N/A	N/A	N/A
Critical Thinking	1,272	154	12%	N/A	N/A	N/A
Multicultural	423	41	10%	80	8	10%
International	31	5	16%	74	9	12%
Writing Intensive	395	55	14%	588	71	12%
Core II	1,195	150	13%	N/A	N/A	N/A
<b>Total</b>	<b>3,942</b>	<b>459</b>	<b>12%</b>	<b>742</b>	<b>88</b>	<b>12%</b>

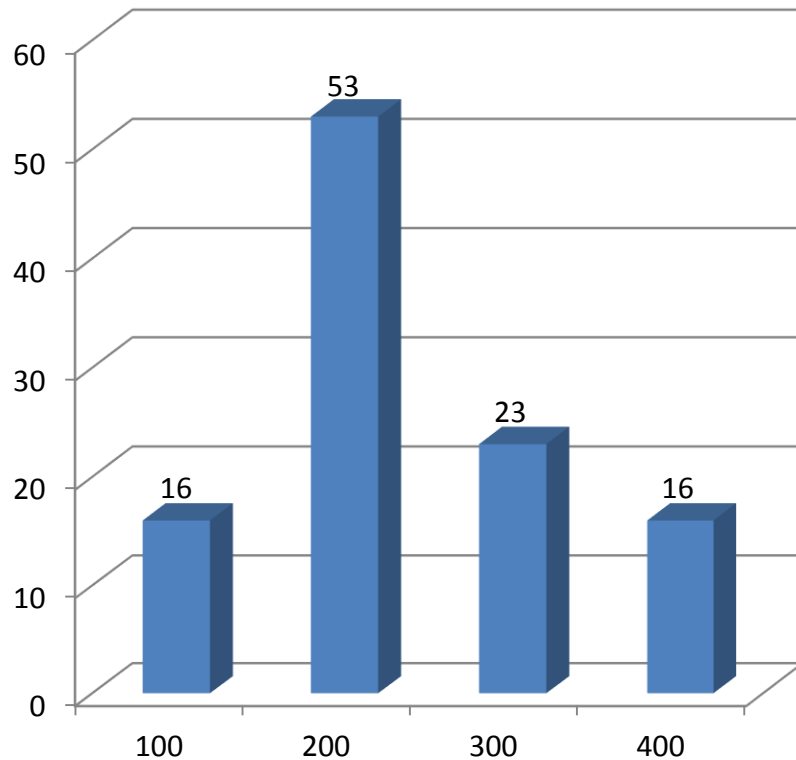
# Population/Sample Comparisons for Marshall's Learning Outcomes by Course Level

Marshall Outcomes	Course Level = 100/200			Course Level = 300/400		
	Population	Sample	Percent	Population	Sample	Percent
Creative Thinking	380	69	18%	202	39	19%
Inquiry-Based Thinking	1,125	82	7%	342	26	8%
Quantitative Thinking	637	93	15%	118	15	13%
<b>Total</b>	<b>2,142</b>	<b>244</b>	<b>11%</b>	<b>662</b>	<b>80</b>	<b>12%</b>

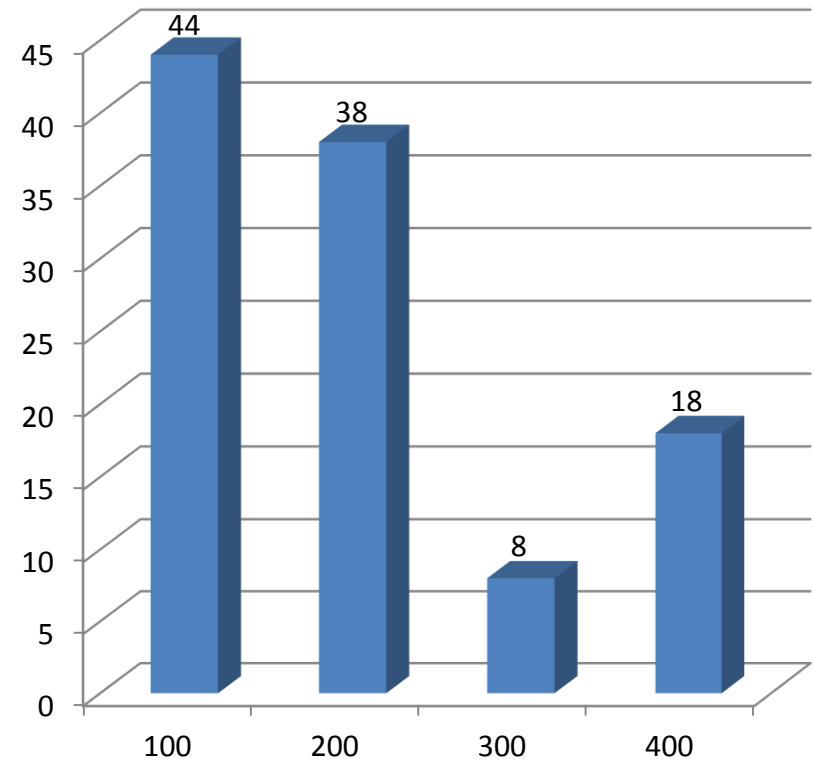
# Sample Frequencies

Total # of artifacts assessed = 108 per outcome

**Course Level Frequencies:  
Creative Thinking**



**Course Level Frequencies:  
Inquiry-Based Thinking**

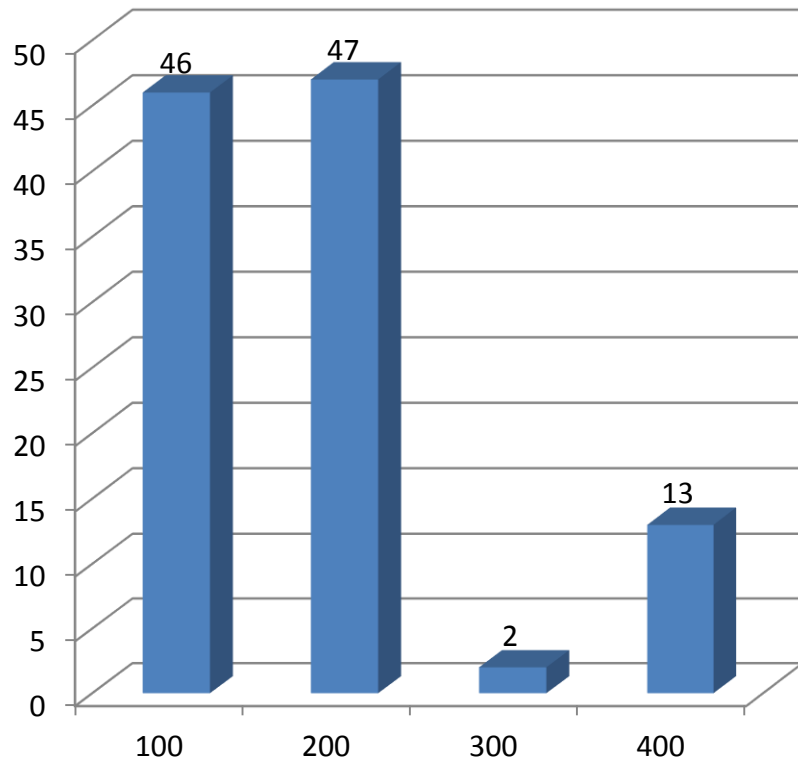


# Sample Frequencies

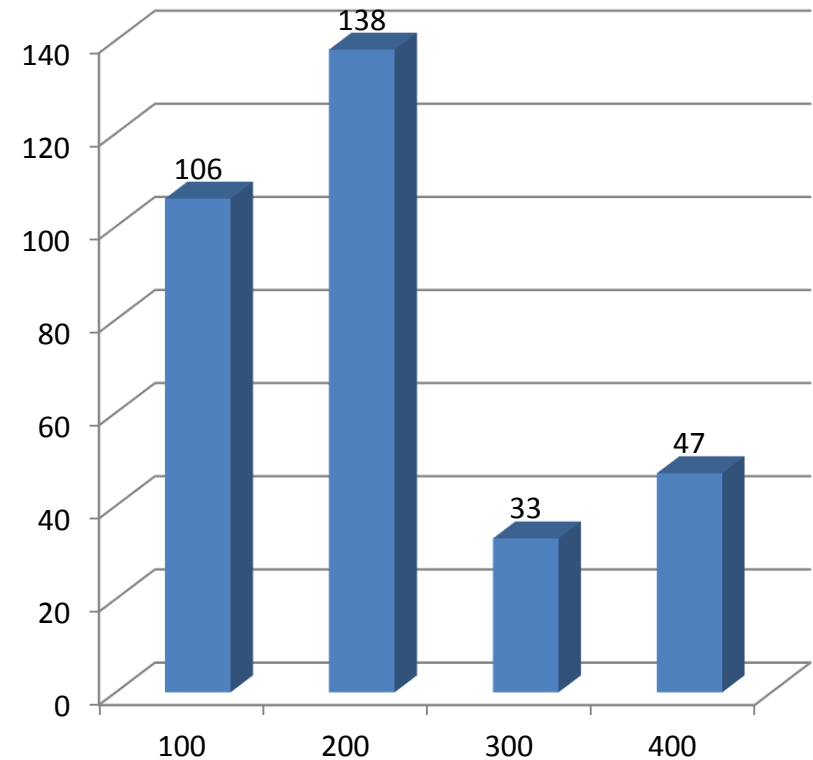
Total # of artifacts assessed = 108 per outcome

Total = 324

**Course Level Frequencies:  
Quantitative Thinking**



**Course Level Frequencies: Total  
across the three outcomes**



# Review Procedures

- Each artifact had two independent raters and usable scores on the 0 – 4 scale were determined in the following manner:
  - If raters assigned the same score, that became the score for the artifact.
  - If raters' scores differed by one point or less, e.g. Rater 1 assigned a score of 1 and Rater 2 a score of 2, the final score was the mean, i.e. 1.5.
  - If raters' scores differed by more than one point, e.g. Rater 1 assigned a score of 1 and Rater 2 a score of 3, the raters met to discuss the rationale for their scores to see if they could agree on a score or, at minimum, scores that differed by no more than one point.
  - If raters' scores differed by more than one point and, after discussion, they were not able to resolve the differences, a third rater was assigned to review the artifact.

# Review Procedures

- We also allowed reviewers to assign a score of *N/A* (*not applicable*) when they judged the assignment to not be aligned with the outcome trait, or a score of *Error* if there was a student upload error or other technical issue which prevented the reviewers from scoring the artifact. When one rater assigned a score of *N/A* or *Error* and the second rater assigned a score of 0 – 4, they also met to discuss the rationale for their scores to see if they could agree on a score or, at a minimum, scores on the 0 – 4 scale that differed by not more than one point. If they could not agree, a third reader was assigned.



## Third Readers for this Year's Review

We had twelve artifacts (with a total of fourteen traits) that required third reviews. In each case, the disagreement was between a score of either a score of *N/A* and a score on the 0 – 4 scale or between two scores on the 0 – 4 scale. All but four of the traits (which were eliminated from the analysis) were able to be reconciled with the third reader.

## Artifacts Excluded from Analysis of Means Due to Inability to Assess or Misalignment with Tagged Outcomes

Outcome	Total Artifacts	Total Artifacts Eliminated Due to Error	Total Artifacts Eliminated due to Misalignment	Total Used for Analysis
Creative Thinking	108	3	24	81
Inquiry-Based Thinking	108	6	13	89
Quantitative Thinking	108	3	16	89
Total	324	12	53	259

# Revised Creative Thinking Rubric

**Creative Thinking:** Students will **outline** multiple divergent solutions to a problem, **explore** and **develop** risky or controversial ideas, and **synthesize** ideas/expertise to **generate** innovations.

<b>Traits:</b> Performance Indicators/ <b>Performance Levels</b>	<b>Level 0</b>	<b>Level 1</b>	<b>Level 2</b>	<b>Level 3</b>	<b>Level 4</b>
<b>Ambiguities &amp; Possibilities:</b> <b>Outlines</b> multiple divergent solutions to a problem.	Does not <b>outline</b> solutions to a given problem.	<b>Outlines</b> a single solution to a problem, either feasible or infeasible.	<b>Outlines</b> some solutions, although not all might be divergent and/or feasible. (2 ideas)	<b>Outlines</b> multiple divergent and feasible solutions to a problem. (more than 2 ideas)	<b>Outlines</b> multiple divergent and feasible solutions to a problem and considers the potential pros and cons of each solution.
<b>Risk Taking:</b> <b>Explores</b> and <b>develops</b> risky or controversial ideas.	Does not <b>explore</b> or <b>develop</b> risky or controversial ideas.	<b>Explores</b> , but does not <b>develop</b> risky or controversial ideas.	<b>Explores</b> risky or controversial ideas and <b>develops</b> these ideas, but only in a superficial manner.	<b>Explores</b> risky or controversial ideas, and <b>develops</b> these ideas in some depth.	<b>Explores</b> risky or controversial ideas, and thoroughly <b>develops</b> these ideas.
<b>Innovation:</b> <b>Synthesizes</b> ideas/expertise to <b>generate</b> innovations.	Does not <b>synthesize</b> ideas/expertise or <b>generate</b> innovations.	<b>Demonstrates</b> rudimentary ability to <b>synthesize</b> ideas, but this synthesis does not result in innovations.	<b>Synthesizes</b> similar ideas and methods to <b>generate</b> an innovation.	<b>Synthesizes</b> divergent ideas and methods to <b>generate</b> an innovation.	<b>Synthesizes</b> by <b>elaborating or expanding</b> on divergent ideas and methods to <b>generate</b> an innovation.

# Revised Inquiry-Based Thinking Rubric

**Inquiry-Based Thinking:** Students will **formulate** focused questions and/or hypotheses, **evaluate** existing knowledge, **collect** and **analyze** data, and **draw** justifiable conclusions.

Traits: Performance Indicators/Performance Levels	Level 0	Level 1	Level 2	Level 3	Level 4
<b>Problem/Question:</b> Formulates focused questions and/or hypotheses.	Does not <b>formulate</b> focused questions or hypotheses.	<b>Formulates</b> a question and/or hypothesis, but not one that is necessarily focused or manageable.	<b>Formulates</b> a question and/or hypothesis that is focused and manageable.	<b>Formulates</b> a question and/or hypothesis that is focused and manageable and addresses a potentially significant area of inquiry.	<b>Formulates</b> a focused, and manageable question and/or hypothesis that addresses significant yet less-explored aspects of the topic.
<b>Research of Existing Knowledge:</b> Evaluates existing knowledge.	Does not <b>evaluate</b> existing knowledge.	<b>Evaluates</b> some existing research relevant to the problem/question, but only includes those that support one side of an issue or includes information from some questionable sources.	<b>Evaluates</b> some existing research relevant to the problem/question from reputable sources. The review is balanced but not comprehensive.	<b>Uses</b> reputable sources to <b>conduct</b> a comprehensive evaluation of existing research relevant to the problem/question.	<b>Evaluates and synthesizes</b> in-depth relevant information from reputable sources representing various points of view/approaches.
<b>Data Collection and Analysis:</b> Collects and analyzes data.	Neither <b>collects</b> nor <b>analyzes</b> the data.	<b>Collects</b> but does not <b>analyze</b> the data.	<b>Collects</b> but incompletely <b>analyzes</b> the data.	Thoroughly <b>analyzes</b> the data.	Thoroughly <b>analyzes</b> and <b>synthesizes</b> the data.
<b>Conclusions:</b> Draws justifiable conclusions.	Does not <b>draw</b> conclusions.	Conclusions neither <b>address</b> the question and/or hypothesis nor are supported by the data.	Conclusions either <b>address</b> the question and/or hypothesis or are supported by the data.	Conclusions both <b>address</b> the question and/or hypothesis and are supported by the data.	<b>Fulfills level 3 plus suggests</b> how results might apply to other problems or inform future studies.

# Revised Quantitative Thinking Rubric

**Quantitative Thinking:** Students will **analyze** real-world problems quantitatively, **formulate** plausible estimates, **assess** the validity of visual representations of quantitative information, and **differentiate** valid from questionable statistical conclusions.

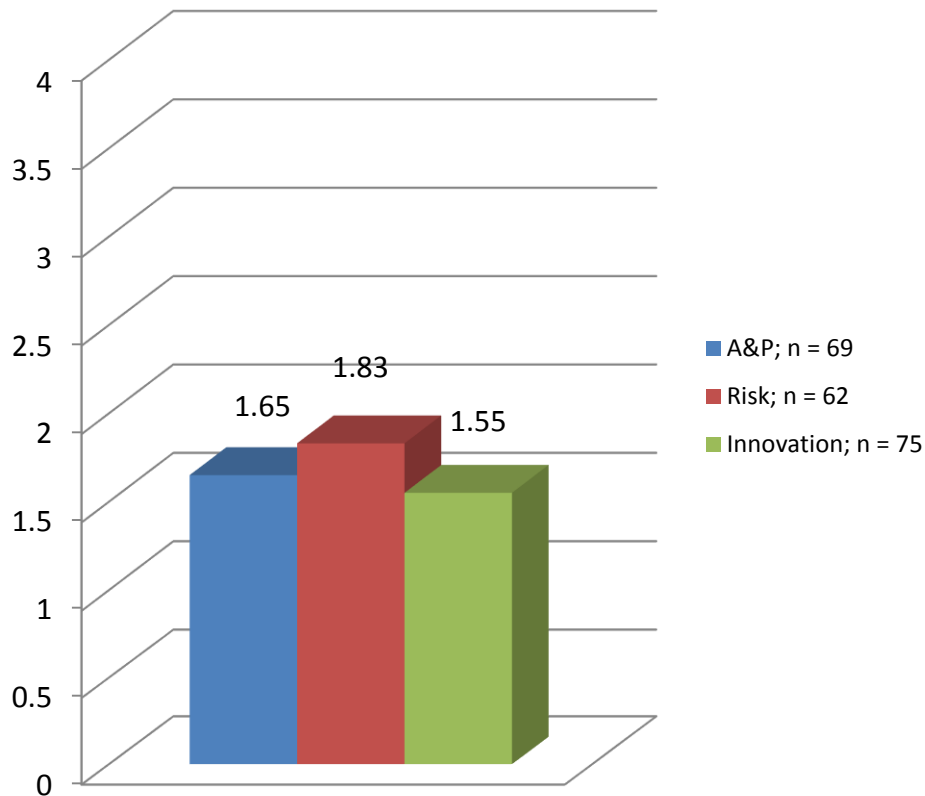
<b>Traits:</b> Performance Indicators/ <b>Performance Levels</b>	<b>Level 0</b> Can't do it	<b>Level 1</b> recognize	<b>Level 2</b> use	<b>Level 3</b> apply	<b>Level 4</b> create
<b>Context:</b> Analyzes real-world problems quantitatively.	Does not <b>explain, report, or analyze</b> real-world problems quantitatively.	<b>Explains and reports</b> the problem within its context quantitatively. <b>Identifies</b> basic metrics to solve the problem.	<b>Level 1 plus uses</b> appropriate tools to analyze metrics to <b>solve</b> problems in a given context.	<b>Level 2 plus articulates</b> meanings of a quantitative analysis.	<b>Develops</b> metrics, <b>uses</b> appropriate tools, and <b>applies</b> solutions to <b>solve</b> novel problems.
<b>Estimation:</b> Formulates plausible estimates.	Does not <b>formulate</b> plausible estimates.	<b>Recognizes</b> applicable units, orders of magnitude and appropriate mathematical tools.	<b>Level 1 plus uses</b> tools to develop the estimate to solve a problem.	<b>Level 2 plus identifies</b> the limitations of the estimated solution as applied to a specified problem.	<b>Level 3 plus Identifies</b> how the purpose of the estimate affects the parameters of the solution.
<b>Visual Representation:</b> Assesses the validity of visual representations of quantitative information.	Does not <b>assess, summarize, or explain</b> the validity of visual representations of quantitative information.	<b>Identifies</b> which visual representation is applicable to construct the model of a problem.	<b>Level 1 plus constructs</b> the visual representation of a model appropriate to the problem.	<b>Level 2 plus recognizes</b> the limitations of how the visual representation of the model applies to the problem.	<b>Level 3 plus identifies, constructs, and recognizes</b> the applicability of the visual representation of the model of the problem.
<b>Statistics:</b> Differentiates valid from questionable statistical conclusions.	Is not able to explain basic statistical terms/concepts.	<b>Recognize</b> types of data appropriate to construct statistical models of the problem.	<b>Level 1 plus uses</b> appropriate statistical tools to <b>estimate</b> models of the problem.	<b>Level 2 plus interprets and applies</b> solutions derived from statistical models of problem.	<b>Independently develops</b> data, <b>uses</b> appropriate tools, <b>interprets</b> and <b>applies</b> results for a problem.

# Creative Thinking

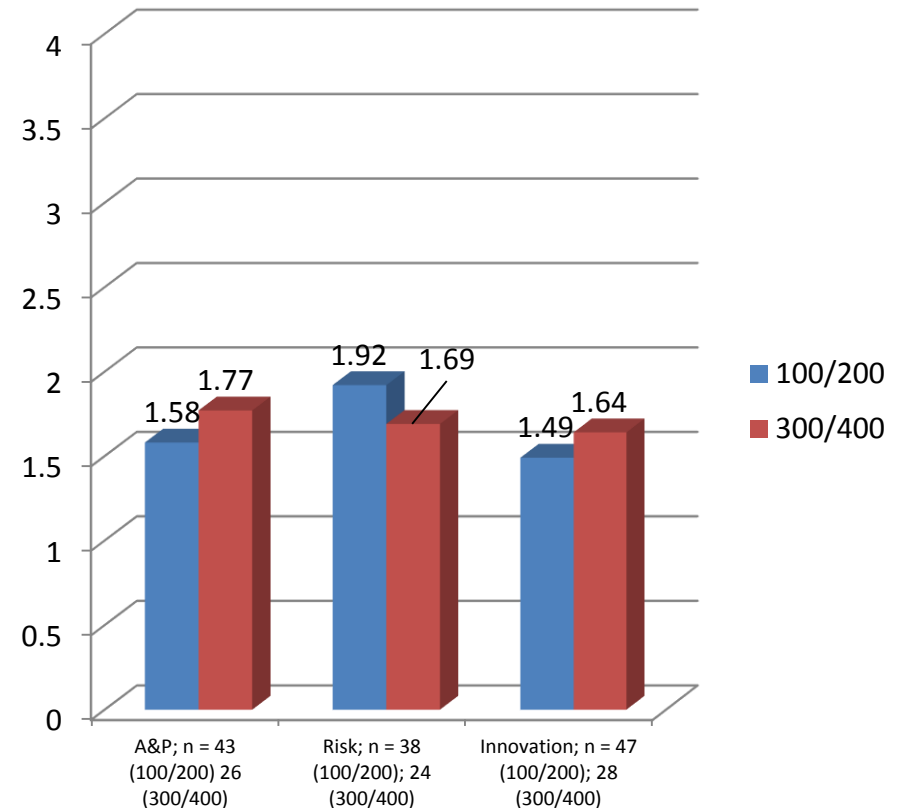
Mean Scores on a scale of 0 – 4, with 4 being the highest possible score  
(Although there were 81 CRT artifacts in the analysis, not all artifacts aligned to every trait)

Mean differences based on course level were not significant.

## Overall Analysis



## Analysis by Course Level

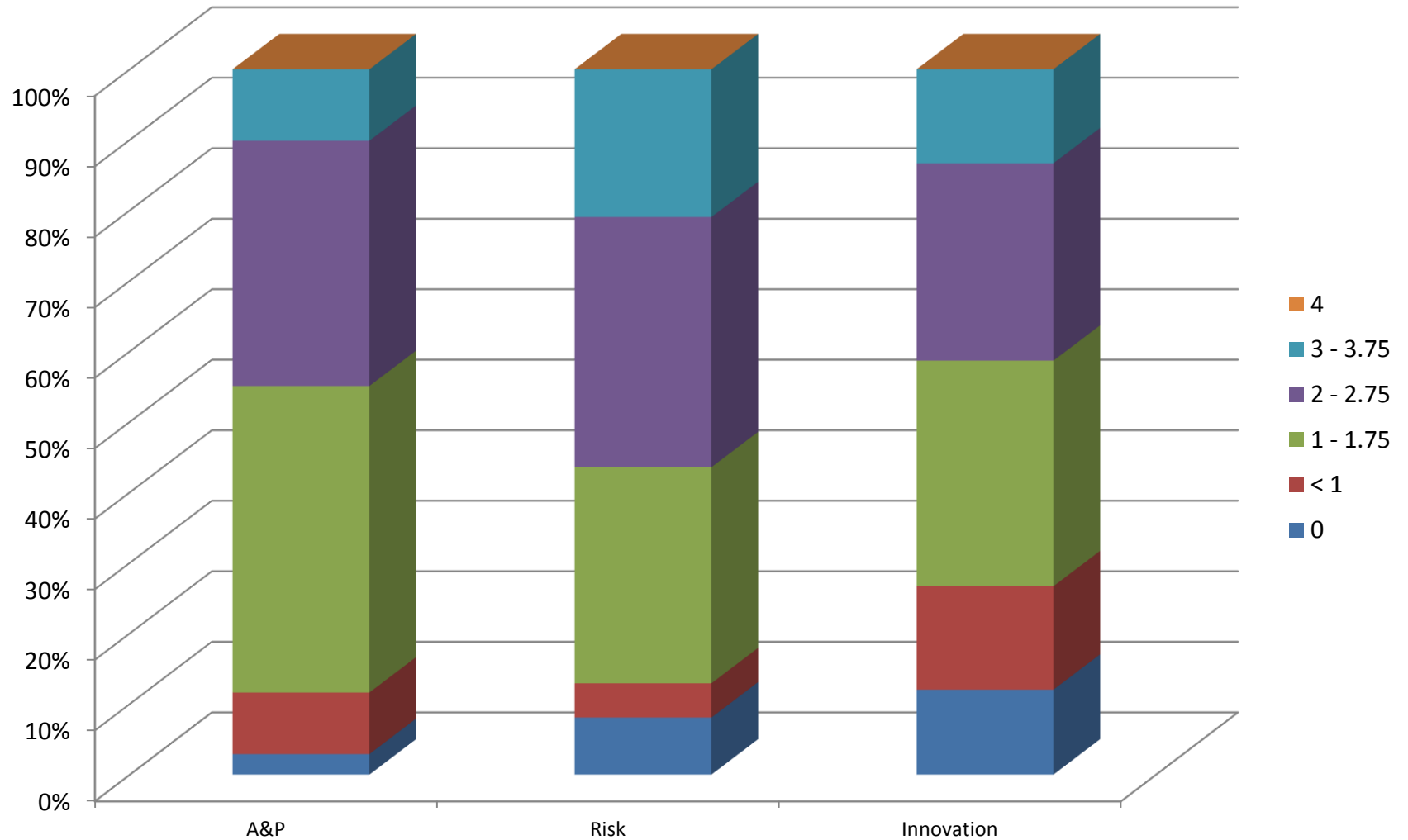


# Creative Thinking

Number of artifacts (with usable scores) scoring at each performance level

Trait/ Performance Level	A&P	Risk	Innovation	Total
0	2 (3%)	5 (8%)	9 (12%)	16 (8%)
> 0, but < 1	6 (9%)	3 (5%)	11 (15%)	20 (10%)
1 – 1.75	30 (43%)	19 (31%)	24 (32%)	73 (35%)
2 – 2.75	24 (35%)	22 (35%)	21 (28%)	67 (33%)
3 – 3.75	7 (10%)	13 (21%)	10 (13%)	30 (15%)
4	0	0	0	0
Total Tags with Usable Scores	69 (100%)	62 (100%)	75 (100%)	206 (100%)

# Creative Thinking





# Creative Thinking

## Inter-Rater Agreement Results

Trait/ Performance Level	A&P; Kappa = .235 (All scores); Kappa = .142 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .574 (All Scores); Kappa Liberal = .865 (Exclusions Noted Above)	Risk; Kappa = .265 (All scores); Kappa = .285 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .466 (All Scores); Kappa Liberal = .782 (Exclusions Noted Above)	Innovation; .092 (All scores); Kappa = .000 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .421 (All Scores); Kappa Liberal = .668 (Exclusions Noted Above)
Agree on score	21 (19%)	21 (19%)	14 (13%)
Difference = 1 point or less	26 (24%)	15 (14%)	26 (24%)
Difference = 1.5 to 2 points	4 (4%)	4 (4%)	16 (15%)
Difference = 2.5 to 3 points	2 (2%)	4 (4%)	0
Difference = 3.5 to 4 points	0	0	0
Agree on Not Aligned	20 (19%)	23 (21%)	11 (10%)
Agree on Unable to Score due to error	3 (3%)	3 (3%)	3 (3%)
Score + Not Aligned	29 (27%)	36 (33%)	37 (34%)
Score + Missing Second Rater Score	3 (3%) (2 of these were missing second rater plus score of not aligned)	2 (2%) (1 of these was missing second rater score plus score of not aligned)	1 (1%)
Total	108 (100%)	108 (100%)	108 (100%)

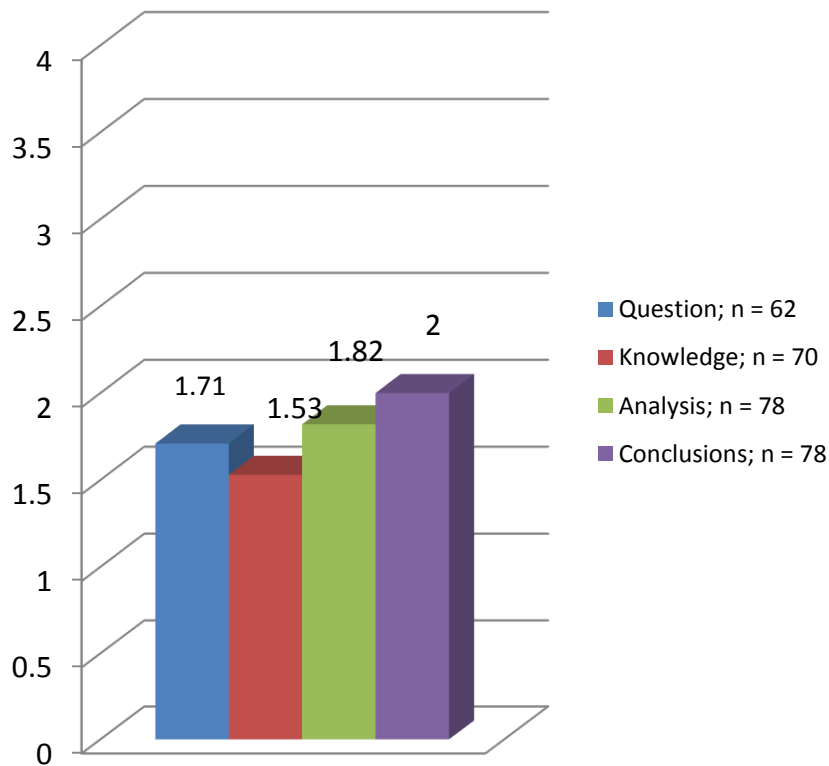
# Inquiry-Based Thinking

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score.

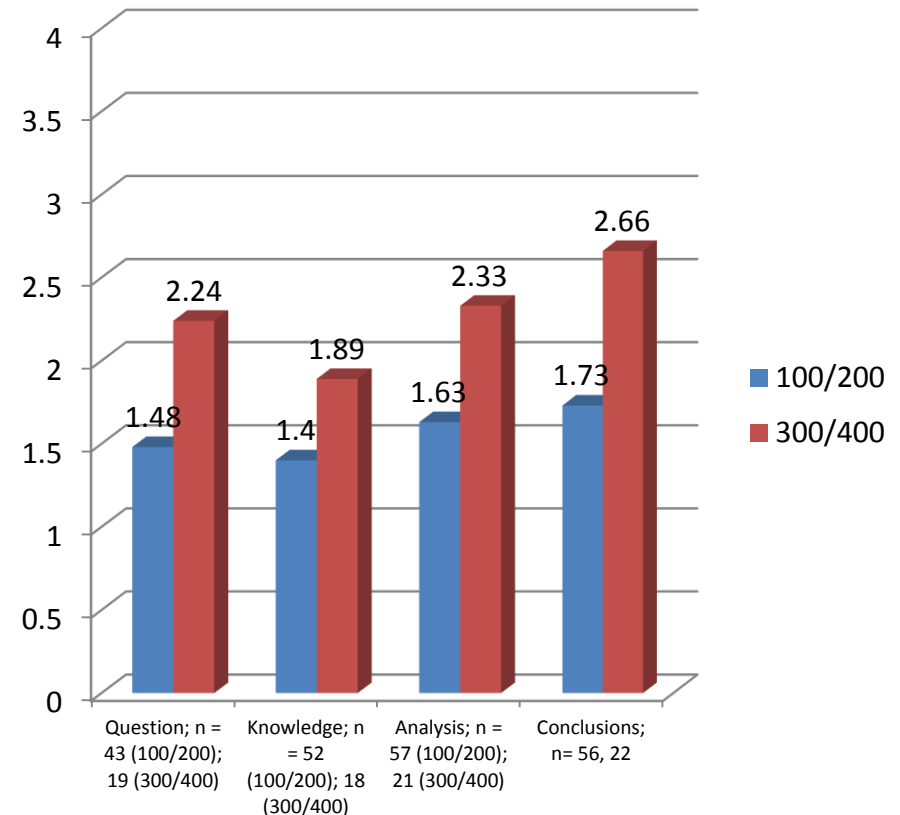
(Although there were 89 IBT artifacts in the analysis, not all artifacts aligned to every trait)

Mean differences for course level were significant for Question/Problem, Data Collection and Analysis, and Conclusions.

## Overall Analysis



## Analysis by Course Level

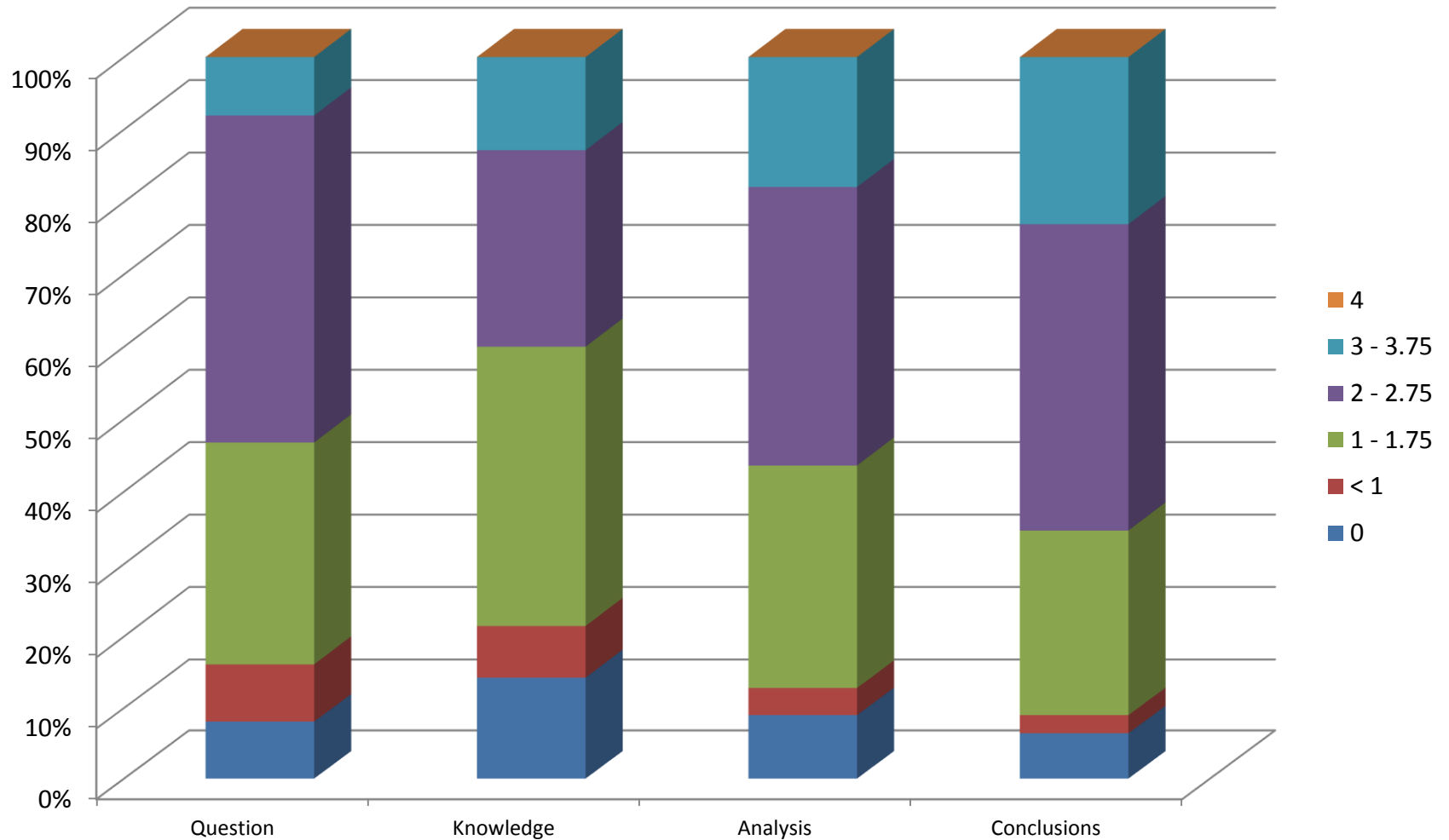


# Inquiry-Based Thinking

Number of artifacts (with usable scores) scoring at each performance level

Trait/ Performance Level	Question	Knowledge	Analysis	Conclusions	Total
0	5 (8%)	10 (14%)	7 (9%)	5 (6%)	27 (9%)
> 0, but < 1	5 (8%)	5 (7%)	3 (4%)	2 (3%)	15 (5%)
1 – 1.75	19 (31%)	27 (39%)	24 (31%)	20 (26%)	90 (31%)
2 – 2.75	28 (45%)	19 (27%)	30 (38%)	33 (42%)	110 (38%)
3 – 3.75	5 (8%)	9 (13%)	14 (18%)	18 (23%)	46 (16%)
4	0	0	0	0	0
Totals	62 (100%)	70 (100%)	78 (100%)	78 (100%)	288 (100%)

# Inquiry-Based Thinking



# Inquiry-Based Thinking

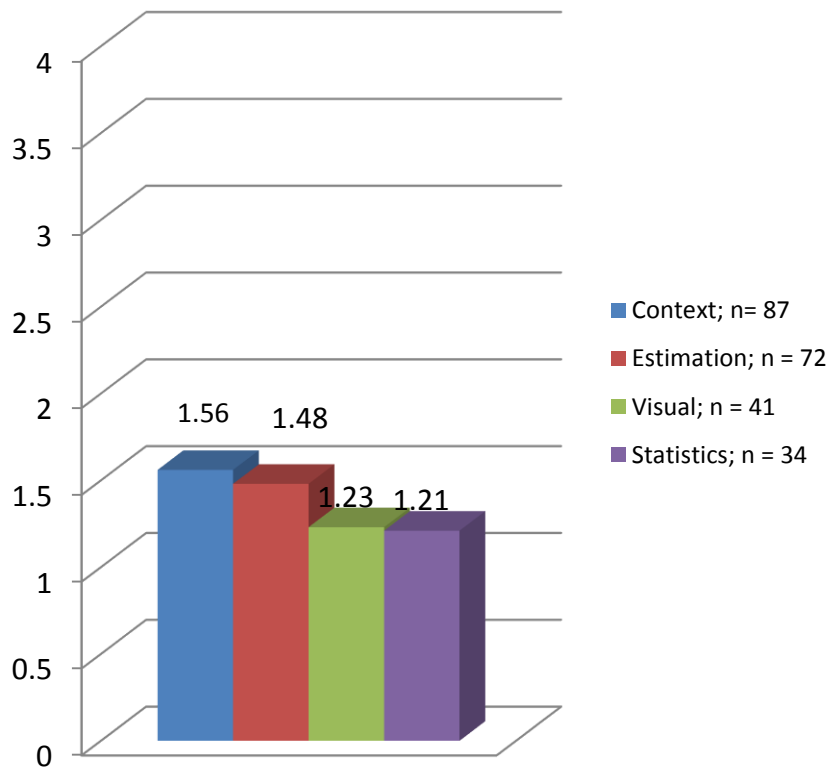
## Inter-Rater Agreement Results

Trait/ Performance Level	Question; Kappa = .263 (All scores); Kappa = .052 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .608 (All Scores); Kappa Liberal = .857 (Exclusions Noted Above)	Knowledge; Kappa = .271 (All scores); Kappa = .161 (Not aligned , unable to score, and one rater score missing excluded); Kappa Liberal = .583 (All Scores); Kappa Liberal = .764 (Exclusions Noted Above)	Analysis; Kappa = .264 (All scores); Kappa = .226 (Not aligned and unable to score excluded); Kappa Liberal = .584 (All Scores); Kappa Liberal = .785 (Exclusions Noted Above)	Conclusions; Kappa = .311 (All scores); Kappa = .222 (Not aligned , unable to score, and one rater score missing excluded); Kappa Liberal = .627 (All Scores); Kappa Liberal = .773(Exclusions Noted Above)
Agree on score	18 (17%)	23 (21%)	29 (27%)	31 (29%)
Difference = 1 point or less	27 (25%)	25 (23%)	25 (25%)	25 (23%)
Difference = 1.5 to 2 points	5 (5%)	9 (8%)	10 (9%)	9 (8%)
Difference = 2.5 to 3 points	1 (1%)	2 (2%)	2 (2%)	4 (4%)
Difference = 3.5 to 4 points	0	1 (1%)	0	0
Agree on Not Aligned	22 (20%)	15 (14%)	9 (8%)	11 (10%)
Agree on Unable to Score due to error	6 (6%)	6 (6%)	6 (6%)	6 (6%)
Score + Not Aligned	27 (25%)	24 (22%)	27 (25%)	18 (17%)
Score + Missing Second Rater Score	2 (2%)	3 (3%)	0	4 (4%)
Total	108 (100%)	108 (100%)	108 (100%)	108 (100%)

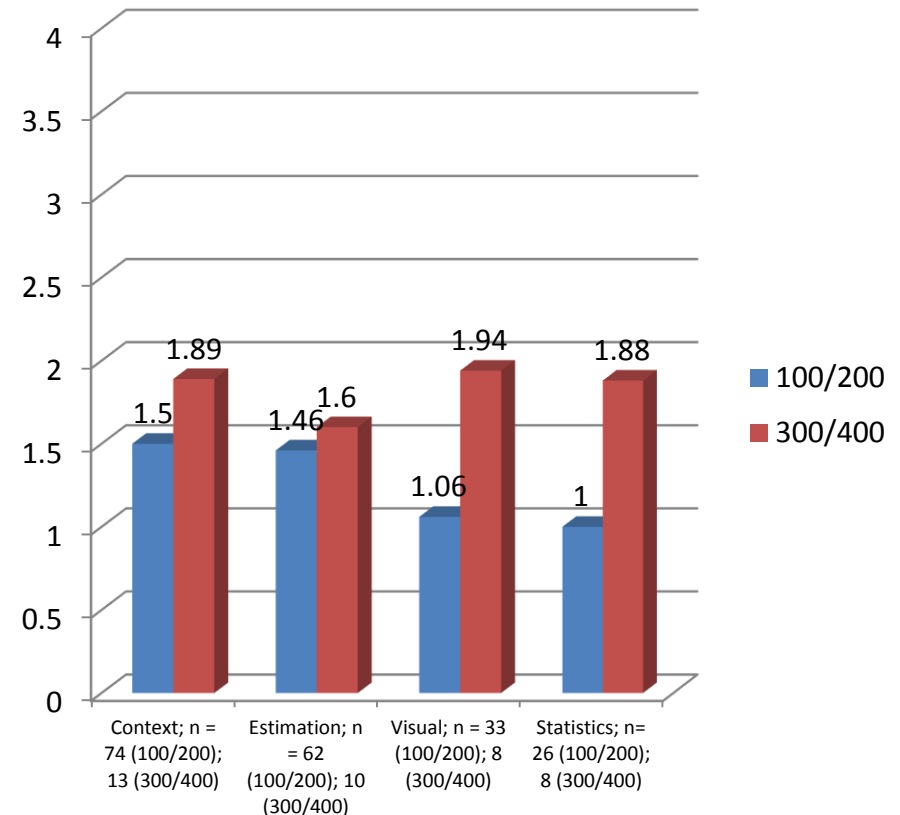
# Quantitative Thinking

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score.  
(Although there were 89 IBT artifacts in the analysis, not all artifacts aligned to every trait)  
Differences were significant for course level for Visual Representations and Statistics.

## Overall Analysis



## Analysis by Course Level

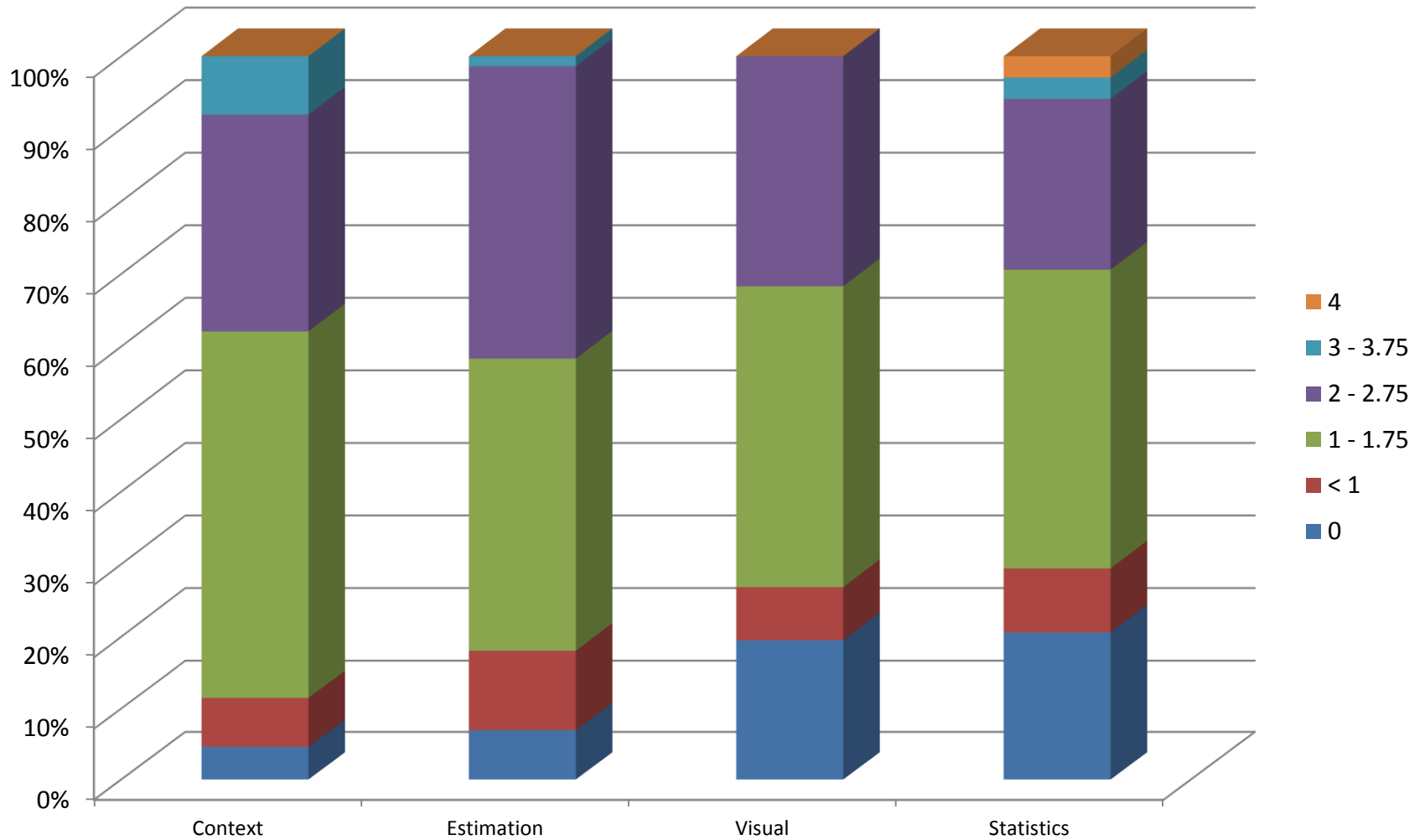


# Quantitative Thinking

Number of artifacts (with usable scores) scoring at each performance level

Trait/ Performance Level	Context	Estimation	Visual	Statistics	Total
0	4 (5%)	5 (7%)	8 (20%)	7 (21%)	24 (10%)
> 0, but < 1	6 (7%)	8 (11%)	3 (7%)	3 (9%)	20 (9%)
1 – 1.75	44 (51%)	29 (40%)	17 (41%)	14 (41%)	104 (44%)
2 – 2.75	26 (30%)	29 (40%)	13 (32%)	8 (24%)	76 (32%)
3 – 3.75	7 (8%)	1 (1%)	0	1 (3%)	9 (4%)
4	0	0	0	1 (3%)	1 (0%)
Totals	87 (100%)	72 (100%)	41 (100%)	34 (100%)	234 (100%)

# Quantitative Thinking





# Quantitative Thinking

## Inter-Rater Agreement Results

Trait/ Performance Level	Context; Kappa = .257 (All scores); Kappa = .190 (Not aligned and Unable to Score Excluded) Kappa Liberal = .645 (All Scores); Kappa Liberal = .805 (Exclusions Noted Above)	Estimation; Kappa = .264 (All Scores); Kappa = .167 (Not Aligned, Unable to Score, and One Rater Score Missing Excluded); Kappa Liberal = .613 (All Scores); Kappa Liberal = .838 (Exclusions Noted Above)	Visual; Kappa = .371 (All scores); Kappa = .170 (Not Aligned, Unable to Score, and One Rater Score Missing Excluded); Kappa Liberal = .585 (All Scores); Kappa Liberal = .873 (Exclusions Noted Above)	Statistics; Kappa = .253 (All scores); Kappa = .344 (Not Aligned and Unable to Score Excluded); Kappa Liberal = .359 (All Scores); Kappa Liberal = .806 (Exclusions Noted Above)
Agree on score	33 (31%)	25 (23%)	12 (11%)	9 (8%)
Difference = 1 point or less	30 (28%)	27 (25%)	14 (13%)	6 (6%)
Difference = 1.5 to 2 points	11 (10%)	7 (6%)	3 (3%)	2 (2%)
Difference = 2.5 to 3 points	0	1 (1%)	0	1 (1%)
Difference = 2.5 to 4 points	1 (1%)	0	0	0
Agree on Not Aligned	9 (8%)	18 (17%)	50 (46%)	50 (46%)
Agree on Unable to Score due to error	3 (3%)	3 (3%)	3 (3%)	3 (3%)
Score + Not Aligned	21 (19%)	26 (24%)	23 (21%)	37 (34%)
Score + Missing Second Rater Score	0	1 (1%)	3 (3%)	0
Total	108 (100%)	108 (100%)	108 (100%)	108 (100%)

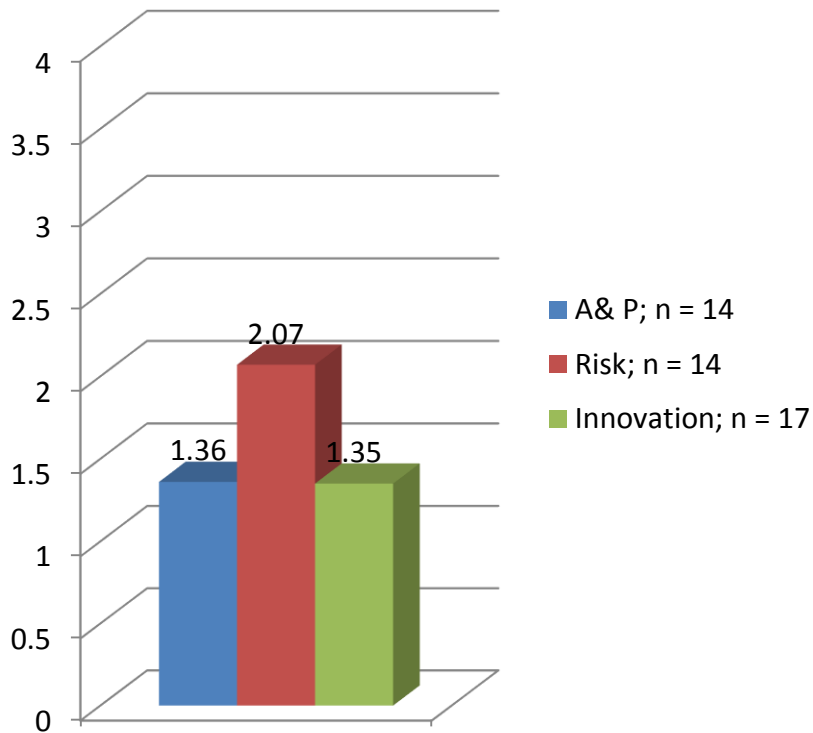


# **Course Type Analysis**

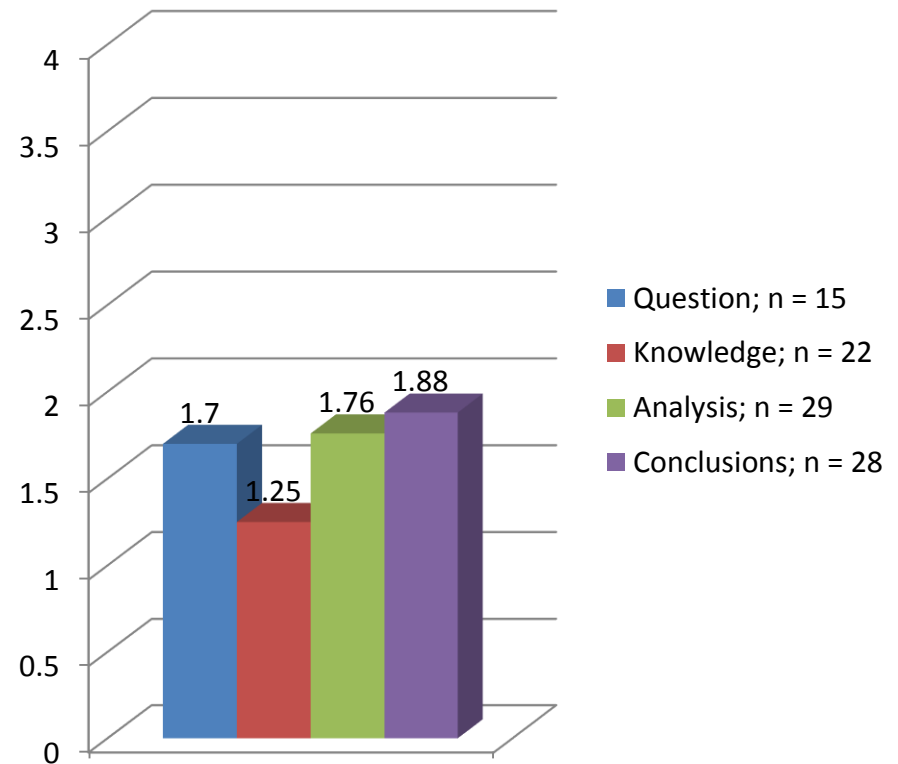
# CT Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All CT courses are 100/200 Level. Some artifacts were from courses that, in addition to being CT, also were multicultural, international, and/or writing intensive.

## Creative Thinking



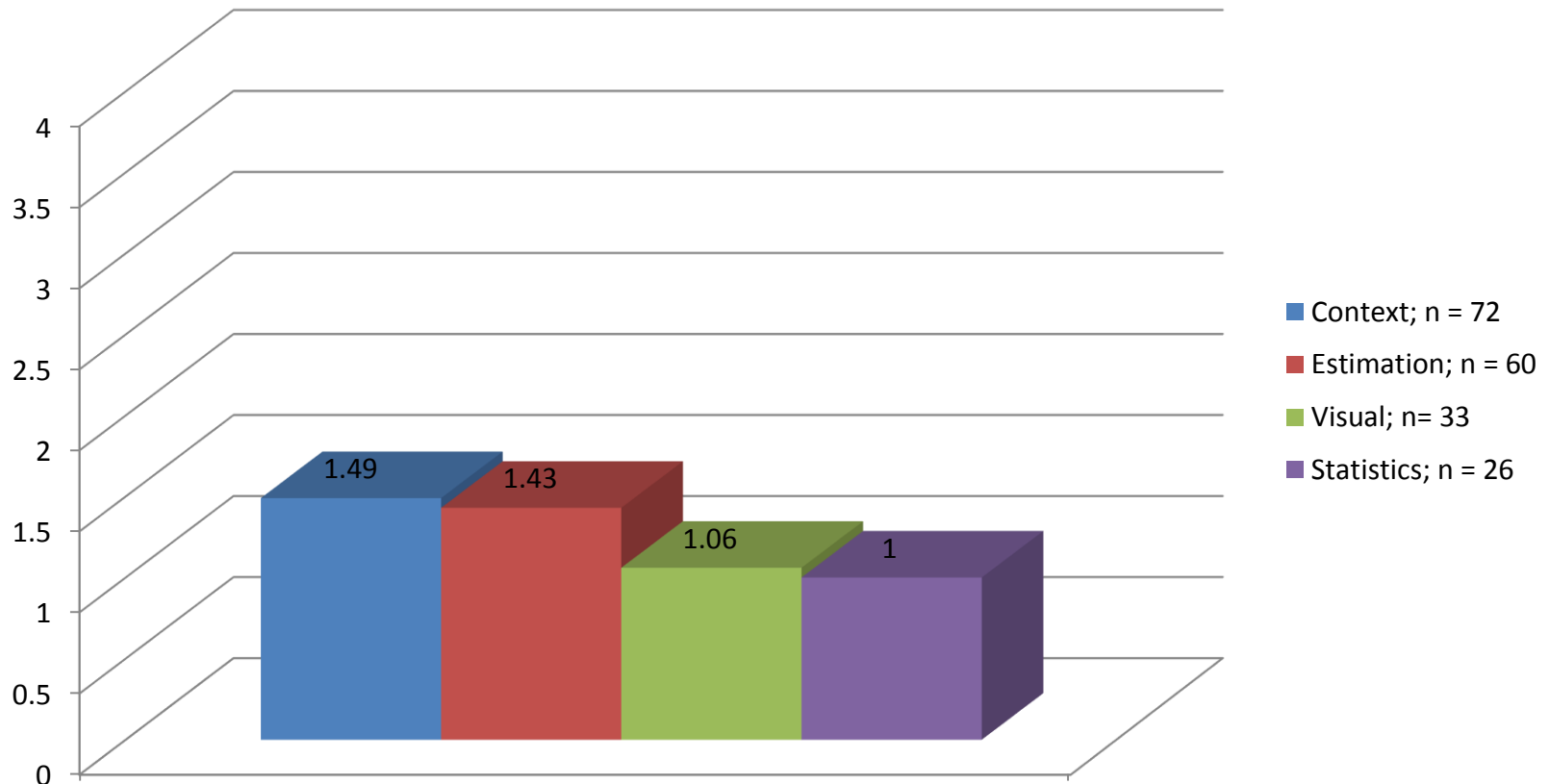
## Inquiry-Based Thinking



# CT Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All CT courses are 100/200 Level. A few artifacts were from courses that, in addition to being CT, also were multicultural, international, and/or or writing intensive.

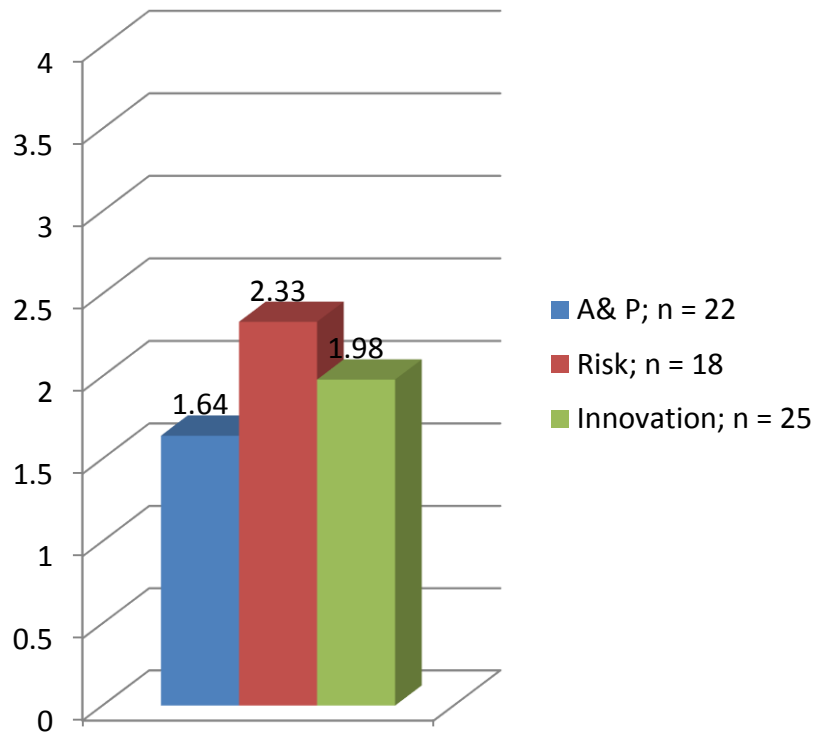
## Quantitative Thinking



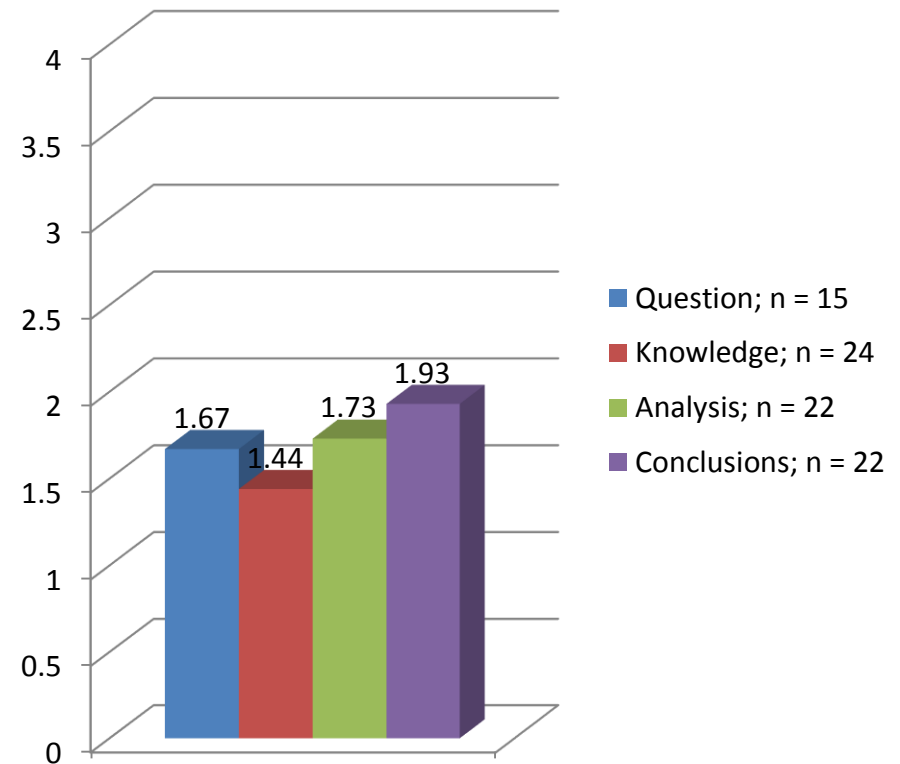
# Core II Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All CT courses are 100/200 Level. Some artifacts were from courses that, in addition to being Core II, also were CT, multicultural, international, and/or writing intensive.

## Creative Thinking



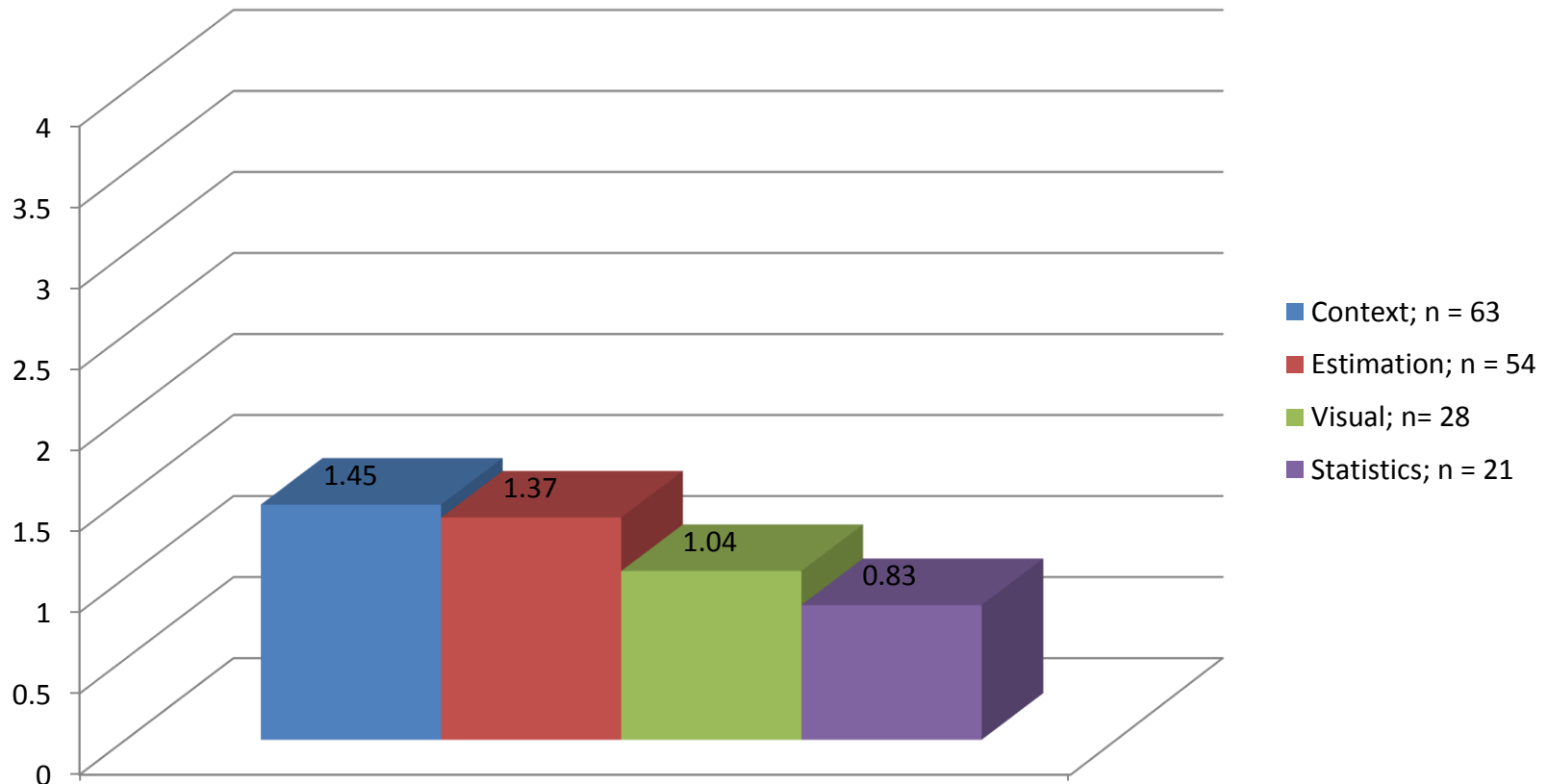
## Inquiry-Based Thinking



# Core II Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All CT courses are 100/200 Level. Some artifacts were from courses that, in addition to being Core II, also were CT, multicultural, international, and/or writing intensive.

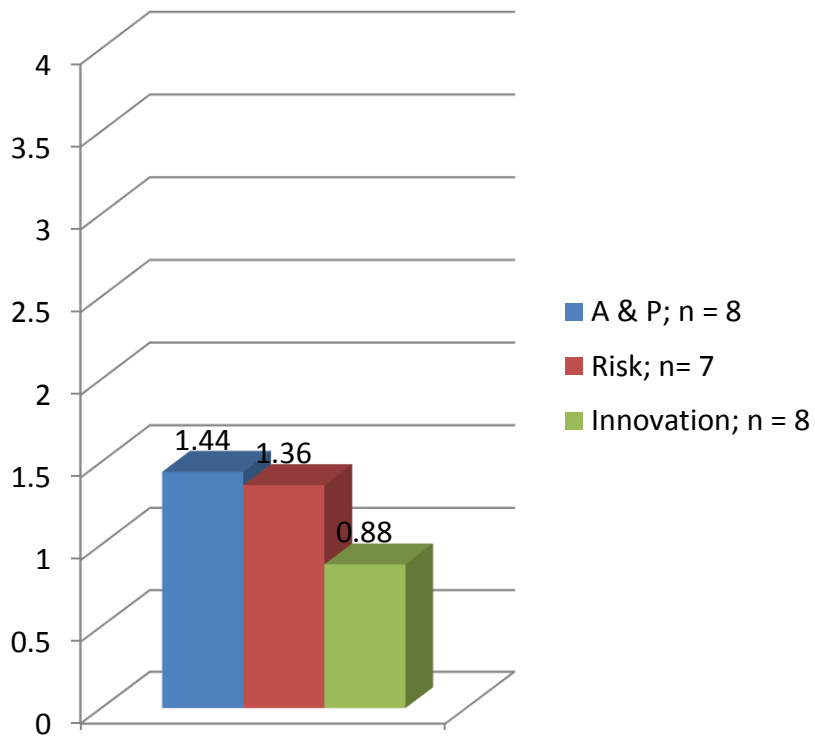
## Quantitative Thinking



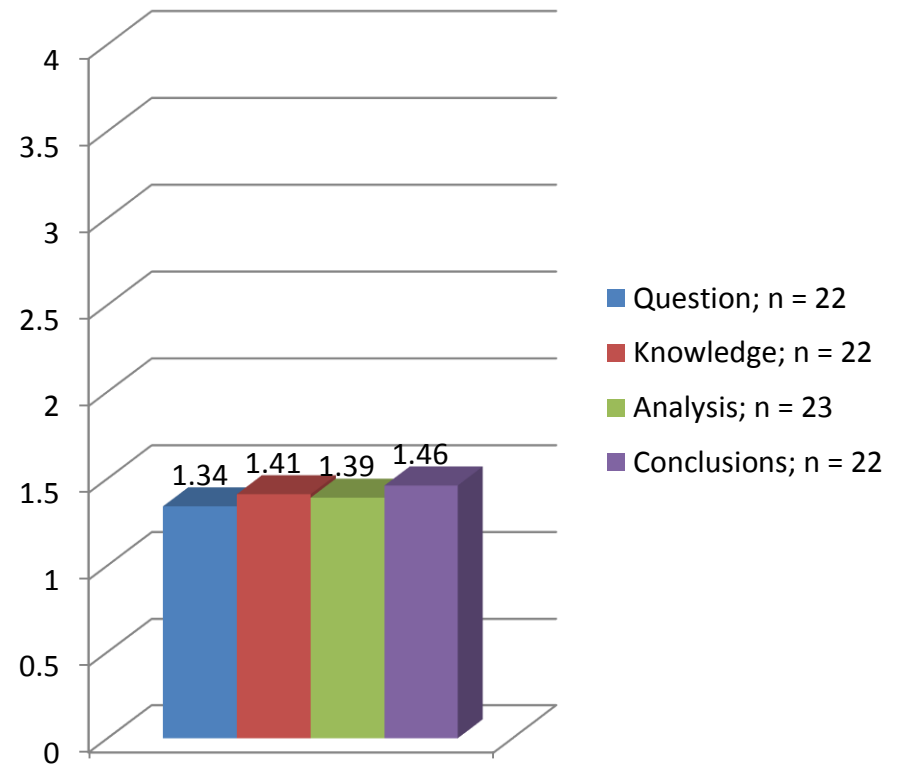
# FYS Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score.

## Creative Thinking



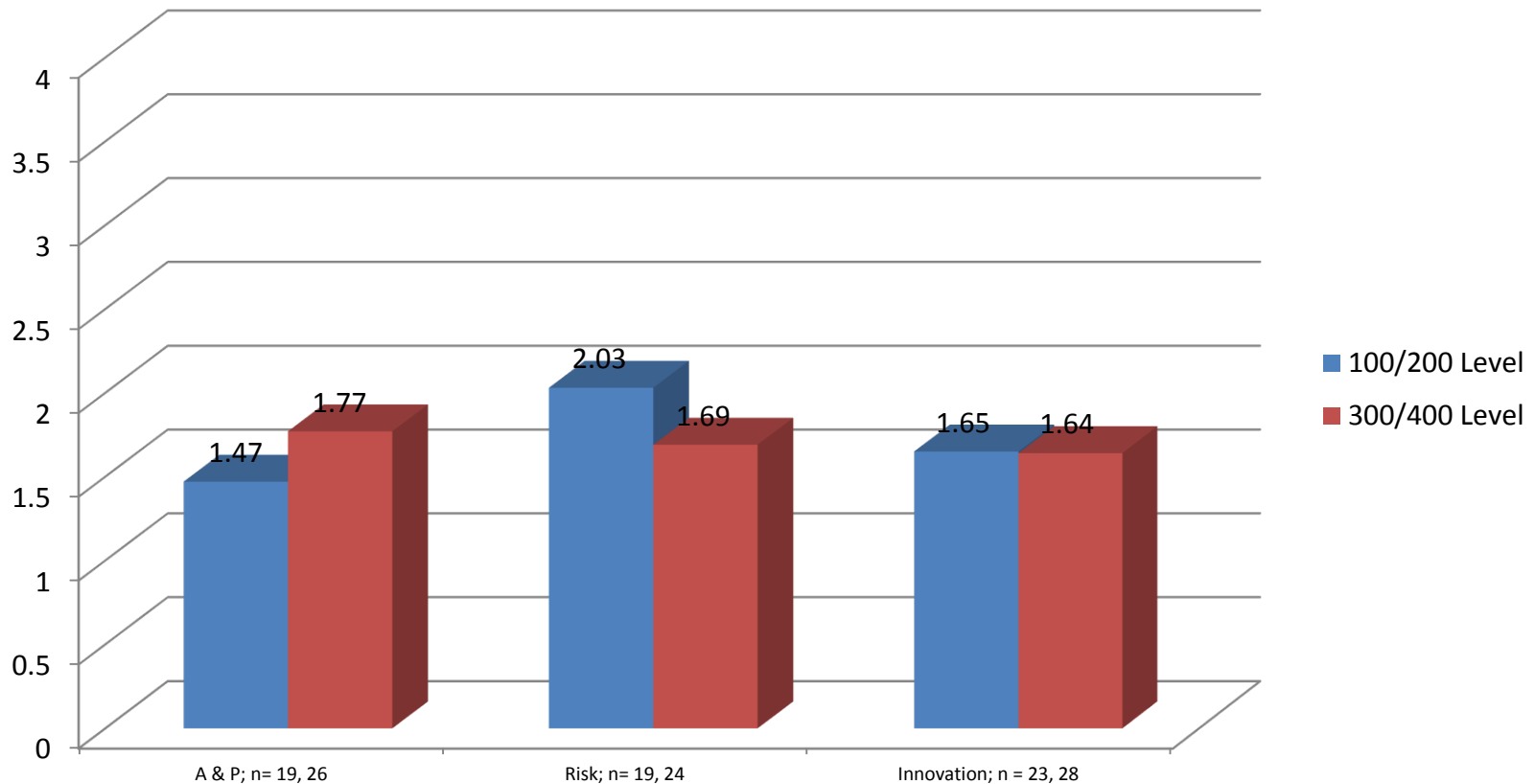
## Inquiry-Based Thinking



# Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, or CT.

## Creative Thinking

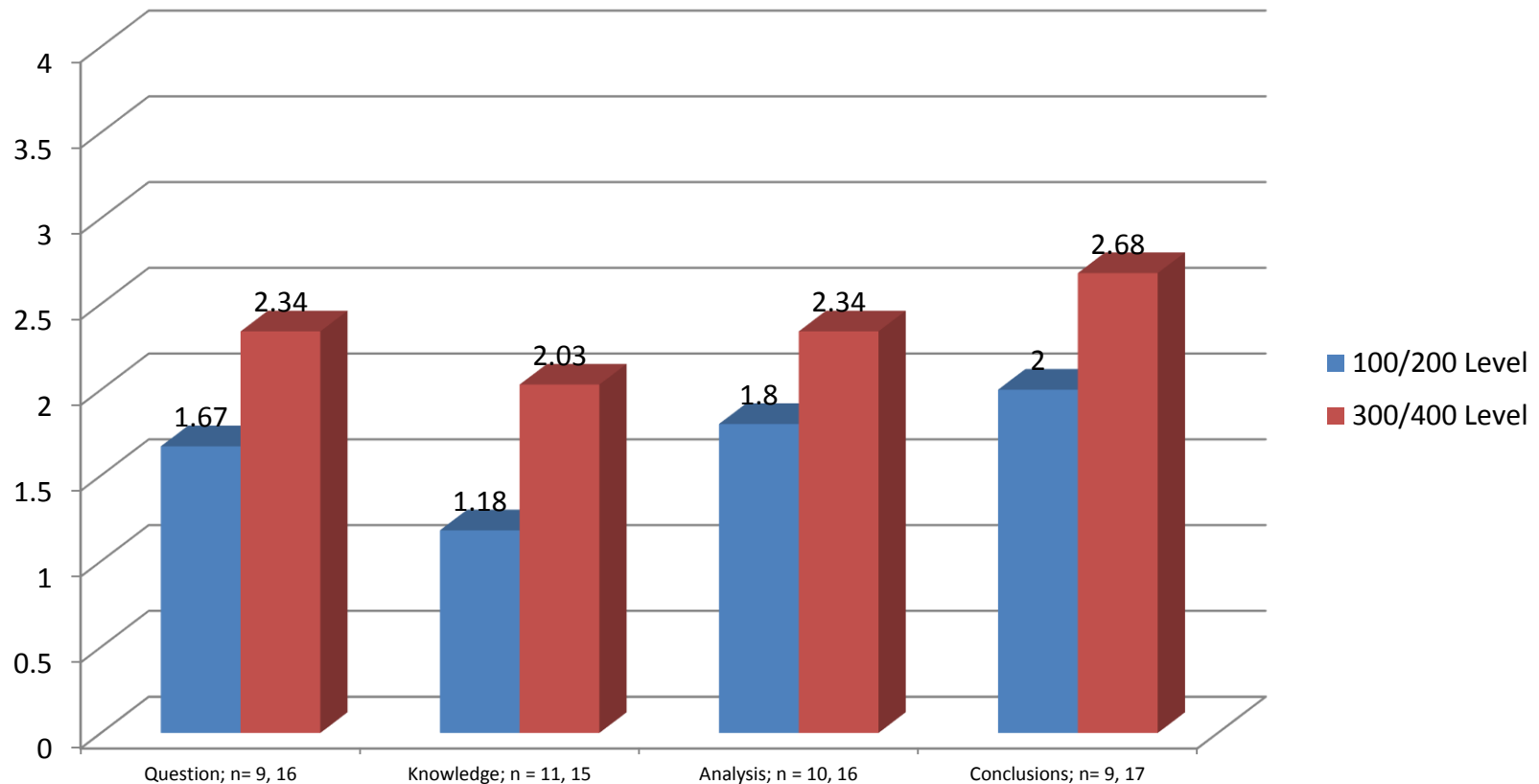




# Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, or CT.

## Inquiry-Based Thinking



# Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, or CT.

## Quantitative Thinking

